Genotype X Environment Interactions in Dairy Cattle.

by

MOHAMED-KHAIR ABDALLA AHMED

A thesis submitted for the degree of
Doctor of Philosophy in the Faculty
of Science

University of Edinburgh
1981

Genotype X Environment Interactions in Dairy Cattle.

by

MOHAMED-KHAIR ABDALLA AHMED

A thesis submitted for the degree of
Doctor of Philosophy in the Faculty
of Science

University of Edinburgh
1981

ABSTRACT OF THESIS (Regulation 6.9)

Name of Candidate MOHAMED-KHAIR ABDALLA AHMED

Address Institute of Animal Genetics, King's Buildings, Edinburgh

Degree Ph.D. *Date* .....

Title of Thesis Genotype X Environment Interactions in Dairy Cattle.

Two sets of heifer records were analysed in an effort to study the problem of genotype X environment interaction in dairy cattle. The first set was of 42000 daughters of 27 widely used bulls. Herds were classified by their average recorded heifer production, the season of calving and herd size. The results of the analysis using the first two classifications indicated the presence of statistically significant sire X class of herd interactions but these were small and were shown to be caused by a small group of bulls which exhibited large deviations over herd classes. Product-moment correlations were high and no significant change in rank was detected although a tendency for increased sire 'variance' with the increase in the level of production was noted. Sire X size of herd interactions were not significant. Interactions involving age effects were studied using regression analysis. Sire X age interactions were significant for milk yield and fat per cent. but the size of the interaction was small.

The second set of data analysed comprised a total of 9775 daughters of 225 young bulls that were tested in four regions. A total of 12610 daughters of 84 old bulls were used in order to create connections between herd-year-seasons. Herds were classified by the mean heifer milk yield and the analysis completed within levels. Yield traits exhibited higher heritabilities in the higher production level while those of percentage traits remained unchanged. Genetic correlations between the performance of daughter groups in the two levels of production were close to one.

It was concluded that there is no need to develop special strains for different levels of production. It was also concluded that there is a significant difference in heritability between levels for milk yield.

Use this side only

I declare that the work contained in this thesis is original and is my own.

List of Contents

List of Tables

List of Figures

Chapter 1

General Introduction and Literature Review

Section 1 - General Introduction

An important objective of most animal breeding policies is the accurate evaluation of the genotypes from their corresponding phenotypes. In practice, the necessary conditions for obtaining unbiased estimates are rarely, if ever, satisfied. Many sources of extraneous variation need to be identified and their effect reduced when possible. However, the evaluation of genotypes would be relatively simple if the relationships between genotypes and environments were just additive. It was recognised a long time ago that apart from the additive effects there are other non-additive relationships (Wright, 1939). Non additivity or non linearity in the relationship between genotypes and environment arises from differential responses of genotypes over environments (Pani and Lasley, 1972). Many methods have been suggested for the classification of genotype X environment interactions. The criteria used in these classifications are: The types of environment, the change in rank of genotypes over environments and the statistical significance of the interaction component of variance.

1) The Theoretical Consequences of genotype X Environment Interactions:

The interaction between genotype and environment can be seen as being made up of two parts; the first is that due to changes in the ranking of genotypes over environments. The

implications of this type of interaction are clear: Wright (1939) suggested that in such a situation a separate genotype should be developed for each environment that is large enough to warrant such an enterprise. The second type of interaction is that due to changes in variance between genotypes over environment. It was this second type that was the subject of many suggestions.

Hammond (1947) suggested that selection in the environment that allows maximum expression of a trait will result in greater genetic progress when the improved strains are reared under less favourable conditions. The opposite view was expressed by Lush (1945) who suggested that selection for a trait should occur in the same environment in which the selected genotypes would eventually perform.

Falconer (1952) suggested that genotype X environment interactions may be expressed as genetic correlations. Performance in two different environments is then regarded to be two different characters with different genetic basis. The genetic correlations for performance of genotypes between environments are assumed to be based on both linkage and pleiotropy and in this respect are similar to genetic correlations between traits in the same genotypes and environments (Bowman, 1972). If we consider only two environments then the rate of genetic improvement in the primary environment i.e. the environment where the selected genotypes are destined to be reared depends on the correlation between the genotypic and phenotypic expression of the character, the genetic standard deviation

and the selection differential, expressed in standard deviation units, in the primary environment. This is known as the 'direct response', whereas if the selection was carried out in a secondary environment and the genotypes were switched to the primary environment to be reared, then the genetic improvement there is known as the 'correlated response'. The correlated response in the primary environment depends on the correlation between the phenotypic expression of the character in the secondary environment and its genotypic expression in the primary environment, the genotypic standard deviation in the primary environment and the intensity of selection.

Under conditions of mass selection, Searle (1965), defined the efficiency of indirect selection relative to direct selection as the ratio of the correlated to the direct response. On simplification this turns out to be:

The relative selection efficiency (RSE(X,Y,y) $= r_g \sqrt{\frac{h^2}{h^2_y}}$ (1.1)

where =

r_g = The genetic correlation

hx^2 = The heritability in the secondary environment

hy^2 = The heritability in the primary environment.

It is evident then that indirect selection i.e. selection in an alternative environment rather than the primary one, assuming equal intensities of selection will result in greater genetic progress when :

$$r_g \sqrt{\frac{h^2 x}{h^2 y}} > 1 \qquad (1.2)$$

Searle (1965) argues that since r_g is less than 1, hx must

exceed hy and that the minimum heritability in the alternative environment is $^{hy}/_{r_g}{}^2$ for indirect selection to be preferable. By substituting 1 in place of hx in (1.2), which is the maximum it can be, it is immediately seen that the minimum value for r_g is hy. However, situations may arise where indirect selection will be preferable even if the inequality in (1.2) does not hold such as the case of selection for female fertility in sheep using testes size.

In dairy cattle selection for milk production is carried out among bulls using progeny testing. Searle (1978) described the relative selection efficiency of progeny testing and the situations where progeny testing coupled with indirect selection results in greater genetic progress. This is important in conditions where bulls are progeny tested in one environment or level of production and their progeny are destined to produce in another. Assuming n progeny per bull and a simple record for each offspring, he states that the correlation between a progeny test using an alternative trait (x) and the additive genotype of the parent is obtained as follows:

$$R\ I_{x,y} = r \frac{nxhx}{4 + hx\ (nx-1)} \qquad (1.3)$$

Where R Ix,y is the correlation between a progeny test using x and the additive genotype y of a parent. A commonly used method for detection of interactions is the application of the analysis of variance techniques. The formula for the composition of the interaction component (from Robertson, 1959) shows clearly that the interaction component is composed of two parts, one due to changes in rank and the other due to change in variance or the 'scale' of the environment:

$$6^2_{GE} = \tfrac{1}{2}(6_{G1} - 6_{G2})^2 + 6_{G1}\,6_{G2}\,(1-r_g) \qquad (1.4)$$

where 6_{G1} and 6_{G2} are the genetic standard deviations in environment 1 and 2 respectively and r_g is the genetic correlation between the performance of genotypes in the two environments. Clearly with equal genetic standard deviations the first part of the expression in(1.4) reduces to zero while with a genetic correlation of unity the second part also becomes zero. However a situation can be envisaged where differences in genetic and phenotypic variance do exist while the heritabilities remain equal. Dickerson (1962) points out that with many environments the mean value of the first quantity in(1.4) is the variance of the genetic scale or V(6Gi). Changes in ranking may or may not be accompanied by changes in genetic variance and in either case the interaction component may or may not be significant. Robertson (1959) suggested that the estimate of the genetic correlation as a quantitative expression of genotype-environment interaction is of value in giving a measure of the biological importance rather than the statistical significance of the interaction. He also pointed out that the normal test of significance for the interaction component seems to be more sensitive to changes in ranking than to changes in variance. He suggested that an estimate of the genetic correlation around 0.8 would be of biological or agricultural importance and that no experiment on interaction would be considered worth doing unless a genetic correlation of 0.6 (significant difference from unity) can be detected.

Under the assumption of similar genetic variance in each

of the environments and a random model, Dickerson (1962) gave the following formula for the genetic correlation (r_g):

$$r_g = \frac{6G^2}{6^2_G + 6^2_{GE}} \qquad (1.5)$$

Clearly inequality of variance would bias 6^2_{GE} by an amount equivalent to the change in variance. On correction for this bias the formula becomes:

$$r_g = \frac{6^2_G}{6^2_G + 6^2_{GE} - V(6G)} \qquad (1.6)$$

where V(6G) The variance of genetic standard deviations

6^2_{GE} The interaction component of variance.

Pirchner (1968) points out that genotype-environment interaction due to changes in variance is generally related to the scale of the environment and can be removed by expressing performance as a deviation from the respective environment mean or by appropriate transformation of the data. The effect of this type of interaction is to change heritability and selection differential and so will change the effectiveness of selection and ultimately the breeding plans. Yamada (1962) gave corrected formulae for the genetic correlation using models other than the random model assumed by Dickerson (1962).

2) Importance of Genotype X Environment Interactions in dairy cattle

In Western Europe and North America the environments to which dairy cattle are exposed are fairly uniform. However, there are certain factors such as the degree of managerial competence and level of concentrate feeding that could, possibly, cause important differences between herds. The need to test

bulls which are to be used in herds with varying management and production levels raises the question of where the records necessary for testing bulls should come from. Also the development of A.I. and deep freezing techniques has for the first time presented the possibility of using semen over vastly different geographic and climatic regions. If the ranking of bulls in high level or pedigree herds is different from that in commercial herds or if their rank in their original country is appreciably different from that in the country where the semen is to be used then depending on the extent of the interaction and the costs involved, it would probably be advisable to obtain records for each bull from the environment in which its offspring are destined to produce. In order to obtain greater genetic progress at the same cost a situation may arise where discriminant use of records is worthwhile due to a higher heritability in one environment than the other and hence the prospect of greater genetic improvement by selecting in that particular environment.

Within this project the problem of genotype X environment interactions was tackled using a large set of heifer records provided by the Milk Marketing Board (M.M.B.). Two subsets of data were drawn from this original set. The first was data on widely used bulls which was analysed to obtain significance tests and product moment correlations between least squares means of bulls obtained in different environments. The second subset of data was on young bulls and was analysed in order to obtain heritability and genetic correlations estimates over

different environments and thereby giving an indication as to whether the variance and the ranking of sires changes from one environment to the other.

The herd 'environment' was classified in different ways such as management level and the seasonal mode of calving in the herd. The lack of detailed information on the herds made independent classification difficult. However, the fact that there are very few reports on this subject on British data justifies a detailed look at the problem.

Section 2

General Literature Review

Genotype X Environment Interactions in Dairy Cattle

There is a large body of information on the problem of genotype X environment interaction in dairy cattle. However, the majority of reports relate to North American populations of Holstein-Friesians. The relevance of these results to the British Friesian population cannot be objectively assessed with the present level of knowledge of the differences between the two populations. However, owing to their similar historical origin and the similarities of the environments under which they produce, the North American results should be of value in specifying the main problem areas and the appropriate methods for tackling the problem.

A major feature of the literature reviewed below is that in most instances it seems clear that authors have failed to obtain detailed information on management practices and most have settled for a rather vague definition of environment such as production levels that are used as an indication of management level. However there are some reports where the 'herd environment' or 'management level' are defined more precisely in terms of feeding practices, housing systems and other criteria. From examining the literature it seems reasonable to assume that whenever the environment is specifically defined the chances of picking important interactions are improved. Some authors have indicated that estimates of sires' interactions with management particulars are small when

taken singularly but considered together they amount to a significant proportion of the total variance.

The methods used in the reviewed literature vary widely; the most commonly used method for detecting interactions is the statistical significance test (F. test) for the interaction mean square term. Some authors have used a full model that accounted for genotype X environment interaction and a reduced model that did not take account of it and the genetic parameters estimates obtained in both ways are then compared. Genetic correlations are often used to detect changes in rank and variance over environments. Estimates of heritability obtained at different levels of production or different environments have been used to detect differences in the effectiveness of selection.

In what follows some of the literature concerning genotype X environment interactions will be reviewed and the emphasis will mainly be on milk yield. There is less information on fat and protein yield and percentage.

1. Sire X Herd Interactions:

There is considerable disagreement in the literature about the magnitude and importance of sire X herd interactions. They have been variously attributed to feeding levels, milking techniques, preferential treatment of the offspring of particular sires, non random mating and a number of other factors. Many of the reports show that sire X herd interactions are not important for the major dairy characters. This is not surprising because of the uniformity of the environment to which dairy cattle are exposed within any one geographic region and the individual attention that has to be paid to dairy cattle.

Kelleher, Freeman and Lush (1967) did not find evidence of the existence of any real sire X herd-year-season interactions. The interaction component was 2.0% of the total variance in milk yield. This estimate was obtained from an orthogonal subset of data comprising nearly 3000 records. However, when the whole data were analysed (more than 37000 records) the estimate of the interaction component rose to 4.6% of the total variance in milk yield. It was concluded that the two estimates were not different and that the latter estimate could have resulted from the non orthogonality of the data. These results are in agreement with the findings of Wadell and McGilliard (1959), Legates et al (1956), Touchberry et al (1960), Hickman and Henderson (1955) and a number of others. However it should be mentioned that the genotypes and environments described in these reports varied widely.

There are reports that indicate the presence of important sire X herd interactions. Specht and McGilliard (1960) found that 7% of the total variation in milk yield was due to interactions between herds and sires. Fairchild et al (1963) reported that sire X herd interactions account for 5.5% of the total variance in milk yield in certain types of adjusted records though it was generally low in others.

Allaire and Gaunt (1965) found that in an analysis of milk and fat yield records which were not adjusted for environmental effects the sire by herd interactions accounted for 2.9% of the total variance in milk yield. When the data were adjusted, the proportion of sire X herd interactions was

found to range from 7.8% to 16.4%. The authors suggested that this could have been due to some bias introduced by the adjustments or to a large sampling variance or otherwise it would be true genotype X environment interaction uncovered after adjusting the data for contempreneity.

Tong et al (1977) analysed more than 13500 records of milk yield, fat yield and fat percentage together with more than 7500 records of protein yield and percentage. The percentage of the total variance that was attributable to the sire X herd interaction component was as follows:

Milk yield	4.1%
Fat yield	1.1%
Protein yield	0.3%
Fat percent	2.6%
Protein percent	5.6%

The authors suggested that similar treatment of sire groups, compensatory mating of sires with certain strengths to cows of certain weaknesses or true genotype X environment interactions could account for the above estimates. The authors also suggested that as milk is a major source of revenue for the dairy man his attitude towards treatment of high proof sires' daughters and the mating system he employs could generate interactions. These factors could not be ruled out with regard to fat percentage since it is a character of considerable economic importance. However, the authors indicated that true genotype X environment interactions might operate with regard to protein percentage since there

are no economic pressures acting on it apart from what is due to its correlation with fat percentage. Heritability of milk yield obtained from a reduced model that ignored sire X herd interactions was 0.49 but was only 0.36 in the full model. Heritability of fat and protein yield did not differ much between the two models. Heritability of fat percentage was reduced from 0.66 to 0.59 and that of protein percentage from 0.46 to 0.31 when sire X herd interaction was accounted for. Under the reduced model heritabilities exceeded repeatabilities for all traits and the author concluded that ignoring interactions can result in marked overestimation of heritabilities.

The fact that herd X sire interactions have often been found to be insignificant does not rule out the possibility of detecting significant interactions when the concept of 'herd environment' is broken down to management particulars such as housing systems, feeding levels, milking techniques or combinations of these factors. Richardson et al (1971) investigated interactions in production and feed consumption traits of dairy cattle. First lactation daughters of 13 sires were assigned to either an all forage ration or to a forage plus grain ration. Two of the sires were proven in New Zealand under a largely forage programme. The differences in production between the two ration groups were large and reflected the stress of the all forage ration. The all forage group produced 83% as much milk, consumed 9.6% more forage but only 58% as much estimated net energy, and was more efficient than the forage plus grain group. The sire X ration interaction

was significant for milk and fat corrected milk production. However the sires ranked the same in the two groups apart from three sires with high interaction constants which included one of the New Zealandean sires. The authors concluded that the magnitude of genotype X level of nutrition was probably negligable within the range of conditions of commercial dairying in the United States. Mao and Burnside (1969) investigated the importance of interactions between sires and each of ten different herd environments. The records were split into groups according to the extent of use of artificial insemination, price received for milk, source of water on pasture, forage feeding methods in winter and summer, amount of exercise received by cows in winter and herd size. For these classes of herd environments no important herd X sire interactions were found. However, a highly significant interaction between sires and level of grain feeding in summer was found. The interaction component accounted for 17% of the total variance in sire proofs and genetic correlations between proofs of the same sires based on their daughters in herds with different levels of grain feeding in summer ranged from 0.54 to 0.79. No conclusive results were obtained on the sire interaction with level of grain feeding in winter and housing systems.

Experiments with identical and fraternal twins have generally indicated that genotype interactions with feeding levels are unimportant. Syrstad (1977) did not find any significant interactions between genotype and level of concentrate feeding in a study involving 60 pairs of Norwegian Red and White

monozygous twins which were fed two different levels of concentrates. In another experiment using identical and fraternal twins Freeman (1969) investigated the effect of genotype X ration interactions using low and high grain rations. The pair X ration interactions estimated from 30 to 40 monozygous pairs at various ages was found to be zero for first lactation production.

2. Sire X Level of Production Interactions:

Division of herds according to production levels is commonly used to indicate different management levels. The difficulty in making general statements about reports on interactions involving management levels is that the criteria used in defining management levels and the genotypes used in each case vary widely. A common technique is to classify records into a number of production levels on the basis of herd-year averages. Robertson et al (1960) argued that as long as the analysis is based on within herd-year differences, which are expected to be independent of herd-year averages, the method of classification will be adequate and statistically independent. They classified herd-years into three production levels and it was found that the variation between sires in milk yield is higher in the high producing herds but since the variance within sires showed the same trend there were no significant differences between the heritabilities of milk yield in different levels. The authors concluded that the correlation between the true breeding value of bulls at different production levels was very high and that there was no need to provide special strains within breeds specially suited to herds of different production

levels and there was no need either to concentrate the daughters of an artificial insemination bull, on which his evaluation is based, in higher producing herds. Thus records from all herds, irrespective of the level of production, may be used with equal confidence. However the authors indicated that the division of herds according to yield which they made will be only slightly reflected in levels of feeding and that a more realistic approach would be to look at feeding practices.

Mason and Robertson (1956) analysed data from the Red Danish breed and divided the herds into three production levels; low, medium and high. The results were clear cut in showing a significant increase in heritability of milk yield from 0.05 $\pm$ 0.04 in the low herds to 0.22 $\pm$ 0.04 in the high herds. They speculated that poorer control over environmental variation, less attention to the principle of feeding according to production and errors in the paternity of daughters might have contributed to the lower estimates of heritability in low producing herds. The ranking of sires at different production levels was found to be the same. However, the estimates of the correlation between breeding values at different production levels had rather high standard errors (0.3 - 0.4). With regard to butter-fat production it was found that the differences in the heritability between the three production levels were not significant. Change in rank was also considered unimportant. The results of this study indicate that it is advisable to concentrate the bulls' progeny in the better herds while being reasonably sure that bulls tested in such a way will rank the same in less producing herds where improvement

is expected to be slower by a half due to the lower heritability of milk production.

Burnside et al (1961) estimated the heritability of milk yield at seven herd production levels and also calculated rank correlations between sires at four herd production levels. Heritability estimates of 0.24, 0.33, 0.30, 0.32, 0.36, 0.35 and 0.26 were obtained from the lowest to the highest production levels, respectively and when plotted against production levels fitted a quadratic curve ($P \leqslant 0.07$). All rank correlations were significant, ranging from 0.46 to 0.78. Comparison of the mean with the overall standard deviations of the sire components of variance also indicated that sires rank essentially the same at all production levels. Similar results were obtained by McDaniel and Corley (1966) who used data on the progeny of 40 sires divided into four production levels according to herd-mate production. Ranking correlations were computed among the sire progeny group means paired in all combinations of the four herd-mate levels. Correlations among the progeny group means ranged from 0.88 to 0.96 over the different herd-mate levels. Relationships between adjacent levels were generally highest, and values decreased as the difference in herd mate levels increased. The authors concluded that artificial insemination sire progeny groups will rank in about the same order regardless of the herd production level at which they are evaluated.

In conformity with other investigations, Van Vleck (1963) found little if any genotype X environment interaction caused

by ranking of sires in four herd-mate levels of production, similar to the results obtained by Robertson et al (1960), he found that both genetic and environmental variance in milk yield were different from one environmental level to the other; the higher the level the more genetic variability. He also found that the proportion of the total variability which is genetic is lower in the lower levels of environment than in the higher levels. This has a bearing on the answer to the question: In which environment should sires be tested? These results indicated that for a fixed accuracy of selection fewer daughters in the higher yielding herds will be needed to prove a sire than in the lower yielding herds. However the exclusion of some herds from the testing programme would inevitably decrease the number of young bulls that can be tested and consequently appreciably large changes in the heritability over different levels of production seem to be necessary to warrant a change in the breeding policy. No change in heritability or rank of sires was detected with regard to butterfat production.

Burnside and Rennie (1968) obtained estimates of heritability at each of seven herd production levels. A lower estimate of heritability at the lowest herd level was found and this was attributed to the lack of full expression of genetic potential by genetically superior cows in the lower herds due to inadequate feeding. A lower heritability estimate was also found in the highest level and this was thought to be due to a greater proportion of mature cows or different culling procedures in the very high producing herds. The authors also found a general

trend of lower genetic and environmental variance at low production levels and increasing variance as the level of herd production increased. Correlations between contemporary comparisons of sires at four levels were highly significant and approached unity. It was concluded that artificial insemination sires will rank the same over all herd levels although records made at low and high levels of herd production may not provide as much information for ranking sires accurately as do those made in intermediate ones. The authors emphasised the need to investigate the specific environments and to examine a broad spectrum of genotypes and the full range of environments. An increase in heritability and genetic variance with increasing mean herd production was also reported by Van Vleck (1966) and by Hartmann (1968) who tested the significance of the correlation between progeny tests in two levels of herd production and the significance of the interaction component. The results did not indicate the existence of important genotype X environment interactions. However the heritability of milk yield was 0.20 ± 0.12 in the low level and 0.36 ± 0.14 in the high level.

In an analysis on butterfat yield records by Legates (1962) there was evidence that the magnitude of additively genetic variance increased at the higher levels of environment. However he suggested that only extreme changes in heritability of a trait are significant with respect to the optimum breeding structure of a population. The optimum percentage of a dairy population which should be bred to

young bulls in a sire sampling programme would vary little for a particular population if the heritability ranged from 0.15 to 0.30. However the improvement from selection is more sensitive to the heritability and even with a given level of heritability differences in the genetic variance may influence genetic change.

3. Age Effects and their interactions

a) The relationship between age and yield:

The relationship between age and yield was recognised before 1900 (Freeman, 1973). Since then the yield-age relationship has commanded a great deal of interest which resulted in the development of a variety of age correction factors. The most important use of these factors is to compensate for differences in production due to variation in age at calving especially in sire evaluation where relatively young progeny of a sire may be compared with older herdmates or contemporaries. This is particularly important in testing programmes, such as the British improved contemporary comparison, that are dependent on first lactations for sire evaluation since young animals are the ones affected most by age differences (Gravir & Hickman, 1966). The milk yield-age relationship, is according to most authors (Hickman, 1957), a curvilinear one resulting in an increase in milk yield up to an age of about three years followed by a steady decline. Sargent et al (1967) found that the linear, quadratic, and cubic effects of age at freshening on milk yield were significant, accounting for 13.4, 7.8 and 0.5% respectively of the total variance in milk yield. Age was found to be a more important source of variation in milk yield than was month of freshening.

With regard to fat percentage and protein, however, the

picture is less clear cut. Gravir and Hickman (1966) found no significant regression of fat percentage with age in the first, second or third lactation animals. Sargent et al (1967) found that the linear regression of fat percentage on age accounted for 1.8% of the total variance and the quadratic and cubic age terms were not significant with all three terms accounting for, about 1.9% of the total variance. Fat % declined linearly with age by 0.0023% per month increase in age. The linear term for the regression of protein percentage on age was found to be highly significant in the study by Sargent et al (1967). The quadratic term of the regression was just below significance at the 0.05 level of probability. Age was found to account for 5.3% of the variation in protein percentage. There was a decline of 0.0024% per month in protein percentage and the effect of age was found to be more important than that of season of freshening. Gacula et al (1968), detected a tendency for milk constituents to decrease as cows get older and that the regression of protein percentage on age was of little significance. Syrstad (1965) also found that the effect of age on protein content was small.

Until recently age correction factors have been calculated subject to the assumption that the effect of other factors such as season of calving and level of herd production remained constant. However, it has been reported (e.g. Syrstad, 1965) that these effects do not always remain constant and consequently the estimated yield-age relationship effects should vary

according to other factors confounded with it. Some of these factors account for a relatively large proportion of the variation in age at first calving. Lee and Hickman (1972) found that differences between herds account for 35 to 40% of the variation in age at calving. Harville and Henderson (1966) reported that differences between sires account for up to 18% of the variation in age at first calving. In the United States both additive and multiplicative age correction factors were developed, some of which take account of interactions.

b) The effect of differences between herds and between herd levels:

These affect the yield-age relationship because herd and age effects are often confounded and because of the varying effect of age from one level of production to the other. Using regressions on age-corrected herd averages Searle and Henderson (1959) developed additive age correction factors that are related to the herd production level. They found that correction factors for age should be larger for higher producing than lower producing herds. Multiplicative age correction factors have been used in some instances in preference to additive factors because the variance of milk yield is closely related to herd average production and increases in production with age tend to be greater in higher producing herds. Miller (1964) found that the additive increase in milk production from 24 months to maturity was substantially larger in high producing herds than in low

producing herds in all seasons. The multiplicative increase with age was found to vary from season to season but was constant for high, medium and low classes of production. He concluded that the production of the herd can be ignored in adjusting for age if multiplicative factors are used. Hickman (1957) found that the reduction in variance due to fitting a individual, linear, within herd regressions was significantly greater than fitting a single breed regression. The author concluded that the change in yield with age is related to the level of feeding and care and as feeding for production is positively related to feeding for heifer development it is to be expected that the yield-age relationship would be different for herds producing at different levels.

The MMB conducted an analysis on records obtained in 1970-71, 1971-72 (Hinks, P.C.). The age range was 20-37 and a total of 78137 milk records and 77463 fat records from the year 1970-71 were analysed together with 79158 milk records, 78463 fat records and 34000 protein records from the year 1971-72. Ages at calving were split into groups as follows:-

Group	Age at calving
1	20 - 23
2	24 - 25
3	26 - 27
4	28 - 29
5	30 - 31
6	32 - 33
7	34 - 35
8	36 - 37

The data were also split into two herd heifer mean levels of milk production as follows:

	Range	Average
Level 1	600 - 800	735 gal
Level 2	1000 - 1300	1082 gal

The study concludes that there was no multiplicative effect due to age with rising herd average and that the curves for the regression of milk yield on age fitted within the two levels were essentially the same. A similar split of the data into levels of fat yield records with an average difference of 48 lb between the high and low levels resulted in the conclusion that the slopes of the two curves were, again, not dissimilar.

c) Sire X Age interactions

This is an aspect of age-yield relationship that received little consideration in the past. In the cases when it has been considered, various definitions of the 'rate of maturity' have been used. Hickman and Henderson (1955) used the difference between first and second lactation to indicate the rate of maturity. They found that its heritability was one fourth to one third that of first lactation milk and fat production. Harville and Henderson (1966) estimated that differences between sires account for up to 18% of the variation in age at first calving. The results of Hillers and Freeman (1965) indicate the presence of differences between sires in the rate of maturity of their daughters. They concluded that the presence of such differences would mean that a ranking of a group of sires on first lactation could differ from a ranking

on second lactations and that the only records that would allow comparisons of sires to be free from the effect of the rate of maturity would be those of mature daughters. However they conclude that waiting for mature records introduces complications due to selection of cows and results in longer generation interval for sires, and that the errors in selection of sires caused by differences in maturity rate are not large enough to warrant increasing the generation interval.

Hargrove (1974) used a set of five definitions for the rate of maturity based on the change in herd mate deviation yield from first lactation to lactations started one, two, three, four and five years later. The sire effect was found to be highly significant in all analyses except in the case of the fifth definition which is to be expected if differences between sires are expected to diminish as their daughters grow older. Heritabilities of the rate of maturity ranged from 0.085 to 0.108 for fat percent and from 0.099 to 0.115 for milk production and they were approximately one third to one half as large as heritability of first lactation milk yield. The author concluded that there were differences between sire groups that could not be removed by other refinements of age effects correction factors and that as high production is needed early it wasn't advisable to wait for slow maturing animals to reach peak production. He also concluded that as long as there was no negative correlation between rate of maturity and first lactation production more rapid maturity would be expected among dairy females though at a slower rate than the improvement in production.

Barker and Robertson (1966) estimated a genetic correlation between first and second lactation that was significantly different from unity. This was an indicator of the presence of differences between sires in the rate at which yield increased with age to maturity. However they regarded the correlation to be sufficiently high that selection decisions can be made initially on basis of the first lactation while a final decision would require the use of second lactation yield of daughters and the increase in yield from the first to the second lactation. They found that the increase in yield from first to second and from first to third lactation had a heritability that was significantly different from zero in a sample of British bulls' daughters. A negative correlation that was not significantly different from zero was found between first lactation production and the increase from the first to the second lactation. If this estimate is to be accepted it would indicate a positive relationship between selection on first lactation and the rate of maturity. The implication of this would be to expect a gradual increase in the rate of maturity as a result of the present practice of selection on the basis of first lactation milk production.

An important aspect of age interactions that is not covered by the present study is the age heifer size relationship. The expectation is that large heifers tend to be bred below the average age and small heifers are bred above the average age. Hickman and Gravir (1968) found that large heifers show a greater milk yield increase per month than do small heifers. The authors pointed out that this means that average correction factors discredit young large heifers and underestimate the potential of older small heifers (Hickman, 1973).

d) Season X Age Interactions:

The importance of seasonal differences of calving in theyield-age relationship has been underlined by many authors. Gravir and Hickman (1966) found that the linear coefficients of regression of milk yield on age were significantly smaller for fall than for spring in first lactation heifers but that was not found to be generally true of second and third lactation cows. Quadratic regression coefficients were significantly different between seasons for the first and second lactations but not for the third lactation. They conclude that it was desirable to use a separate set of correction factors for each season especially for the first lactation. However they found no significant regression of fat percentage with age for first, second or third lactation nor for all animals taken together. Lee (1974) found that a single quadratic regression curve was sufficient to represent the relationship between milk yield and age at calving irrespective of month or season of calving for first lactations. This result is in agreement with the findings of Wunder and McGilliard (1967) who found no evidence of age x season interactions when using three age classifications of two, three and four years or older. Different results were obtained by Mao et al (1972) who analysed a large set of data comprising more than 604400 Holstein records. Age-month of calving constant estimates were obtained by the maximum likelihood method, removing biases due to selection, herd differences and environmental trend. The authors found that monthly age correction factors were sufficiently different from each other, not to be grouped into a few seasons. The

conclusion was that serious bias in adjustments was likely if the effect of age and month of calving were not taken jointly. Similar results were obtained by Miller et al (1970) who found that month of calving by age interaction was of such importance that comparisons of individual records and bull proofs could be badly biased if no adjustments were made for the interaction.

In the MMB study referred to earlier, it was found that the seasonal trend in milk yield was roughly the same over the two years and three broad age groups. The seasonal effects were found to be slightly more pronounced for animals calving between 20-27 months of age. With regard to fat % remarkable inconsistency was found in the seasonal trends between the two years. Although fat percentage showed an upward trend with calving age, the results in the two years did not agree closely. Age effect on fat yield was the same in the two years but the seasonal fluctuations tended to be more pronounced in the year 1971-72. Protein percentage compared to fat percentage had higher seasonal variation. An upward trend with age was detected. Protein yield followed the same pattern as fat yield but seasonal fluctuation was not as pronounced.

e) Other Interactions

Various genotypes, environments and characters were subject of studies on interactions in dairy cattle. Lytton and Legates (1966) studied sire X region interactions over the north and south of the United States. Rank correlations were found to be close to unity and the interaction component of variance was negligable. Interactions of sire X sequence of lactation were

found to be insignificant by Hansson and During (1961) despite a higher heritability in first lactation. There is a large number of conflicting reports on the significance of sire X ration or feeding regimes (e.g. Hancock 1953, Freeman 1969). The conflict may have stemmed from the extent of differences between feeding levels and from the variety of genotypes used. Type of housing, calving interval and days dry were also subject to genotype X environment interaction studies (Burdick and McGilliard , 1963).

Conclusions

To conclude this review it might be useful to point out some of the major features of the literature reviewed above. With regard to all classes of interactions it appears reasonable to assume that when detailed information is available on the environments being studied the chances of picking up important interactions are enhanced. There is considerable disagreement between authors on the importance of various classes of interactions which in many cases may be attributed to the extent and type of differences between the environments being studied and the amount of information that is available on them.

With regard to sire X herd interactions, most reports agree that it is either small or non-existent (e.g. Kelleher et al, 1967, Van Vleck & Henderson, 1961) although there are some reports to the contrary when the 'herd environment' is defined more specifically in terms of feeding and housing practices (e.g. Mao and Burnside, 1969). However the interest in herd X sire interactions which, according to most reports is between 2-3% of the total variance, appears to be purely academic since

it is difficult to see what can be done about it unless the cause of interactions is specifically identified and the herds are classified accordingly into large groups which can be treated separately in a breeding policy.

In reports on sire X level of production interactions there is disagreement on the criteria used to define the herd level of production although in most cases the herd-year mean production was used. There seems to be general agreement that the correlations between various levels of production are generally high thereby eliminating the need for producing special strains for different classes of herds. However some rank correlation which were apparently low (e.g. 0.46 to 0.78 in the study by Burnside et al, 1961) were found to be significant which indicates that they were tested against a zero correlation. Many reports indicate that although the rank correlations were high, the heritabilities were different between the various production levels. Mason and Robertson (1956) reported an increase of 0.05 ± 0.04 in the low herds to 0.22 ± 0.04 in the higher levels. In order to judge whether the difference in heritability warrants a change in breeding policy the magnitude of the difference has to be ascertained and even then it appears that it needs to be a large difference to warrant such a change. The nature of the change will be concerned with the amount of use made of records obtained from different classes of herds and to concentrate on using records from herd classes that exhibit higher heritabilities.

Sire X age interactions were generally found to be significant and it seems reasonable to assume that differences between

sires in the rate of maturity of their daughters do exist (Hargrove, 1974, Hillers and Freeman, 1965). However, there is disagreement as to the size of these differences. There are also practical complications arising from the course of action to be taken assuming that these differences are significant. A possible course of action would to be wait for mature records of cows which results in considerable lengthening of the generation interval and also brings with it the problems of cow selection. Age X level of production interactions, which have been reported to be significant, can be dealt with by developing additive age correction factors that are related to herd level of production in the manner of Searle and Henderson (1959). Other authors (e.g. Miller, 1964) have found that the herd level of production can be ignored when multiplicative correction factors are used.

The above conclusions apply to milk production, fat percentage and protein percentage although in most of the literature reviewed the stress has been on milk production. There is very little information available on either fat or protein production.

CHAPTER 2

Analysis of Records of Daughters of Wide-Spread-Use Bulls

Section 1.

Introduction

In this study a set of 42701 milk records of daughters of twenty seven widely used sires was chosen for the following reasons:

a) The set is of value in answering the practical question as to whether genotype X environment interactions are of importance among widely used sires which contribute a large proportion of the dairy herd. The answer to this question will give an indication to the answer of the more general question relating to the presence or absence of important genotype X environment among the dairy cattle population in Britain.

b) Because of the large sire groups it is possible to obtain reliable estimates of certain classes of genotype X environment interactions e.g. sires X level of herd production and within sire regressions on age at calving. The progeny groups have a minimum of just over 600 and a maximum of over 3600.

However, this set of data also has its limitations, the most important of which are:

a) Due to the wide use of the twenty seven sires over a large geographic area, the number of daughters of any particular sire in a single herd-year-season tends to be small. This reduces the reliability of certain estimates, particularly those involving herd-year-seasons interactions with other effects.

b) Because this is a set of highly selected sires it could be argued that any sires that exhibit genotype X environment interactions are likely to have been removed during the process of selection. In effect this means that this set of data may not

be representative and so no extrapolation of results to the general population of dairy cattle can be made. Also because these are highly selected sires estimates of the components of variance are of little value and only tests of significance and product moment correlations can be calculated.

c) Another limitation arises from the nature of the original set data. The lack of independent information on actual management practices and levels of management made it difficult to study the problem in more detail. For the purpose of classicication of herds the only available source was the data itself, although the herd means which were used in this classification are statistically independent of the actual observations.

Section 2

Materials - Original Data

Four tapes of first lactation records covering the recording years 1971-72 to 1974-75 inclusive were obtained from the Milk Marketing Board of England and Wales (MMB). Each recording year extends from the first of December to the end of November. Each tape comprised more than a hundred thousand first lactation records mainly of the British Friesian breed but included a number of other breeds as well which were discarded within this project. The total numbers of records available on milk yield, fat percentage and yield, protein percentage and yield are shown on table (2.1).

Table (2.1)

Numbers of British Friesian Heifer Records

Tape Number	Milk Yield	Fat Yield and Percentage	Protein Yield and Percentage
1	98247	97442	59822
2	107552	106677	67319
3	102846	101935	59678
4	97952	97189	51728

The tapes cover an overlapping time span and many records were found to be repeats of others on other tapes and were consequently removed. Each record contained information on the sire herd-book number, herd-code, age at calving (in months), year and month of calving, length of lactation, milk yield, and for most, fat percentage and yield. Protein percentage and yield information were available on about half the total number of records. Incomplete records with missing information on age at calving or month of calving were discarded.

The Set of Widely Used Sires:

This set of data was chosen from the original data according to the following conditions:

a) The record must have full information on milk yield, fat yield and fat percentage. Due to the large difference in number between milk and fat records on one side and the protein on the other, the latter were analysed separately in a second data set.

b) Each record must have information on both age at calving and month of calving. The age of calving must not be less than 22 months and must not exceed 38 months. This restriction was imposed to approximate the conditions set by the MMB for records to be included in the Improved Contemporary Comparisons analysis (ICC).

c) Any record made in a herd where only one of the twenty seven sires was used was ignored since it will not contribute to the accuracy of estimating the herd-year-season X sire interaction or the sire main effect. However the division of herds into year-seasons still meant that many subclasses are empty. Having satisfied all other conditions each sire to be selected must have a minimum of 500 records.

Table (2.2) shows the names and herd-book numbers of each of these twenty seven sires and their dates of birth. The serial numbers of the sires given in this table will be used to refer to particular sires subsequently. Table (2.3) shows the latest available I.C.C. results for milk yield, fat percent, protein percent, fat yield and protein yield for this set of sires. The wide use of some of these sires which have low and sometimes negative I.C.C. values for milk yield may be explained in terms of subsequent deterioration of I.C.C., high protein or butter-fat proofs. There is also the possibility of desirable conformation characteristics influencing the use of some of these sires.

Preparation of Data for Analysis:

Depending on the date of calving the data were classified into herd-year-season combinations as follows:

Season 1 December - March inclusive

Season 2 April - July "

Season 3 August - November "

The data were also classified according to the following criteria:

a) Herd level of Production:

The herd-years were classified using herd average 305 day first lactation milk production. In obtaining the herd means the following types of records were excluded:

Table (2.2)

Names and Herd-book Numbers of the Wide-Spread Use Sires

Serial Number	Name	Herd-book Number	Date of Birth
1	Okery Mark	260873	8.4.60
2	Ringswood Leader	280963	18.8.62
3	Calcourt Magnus	280983	10.9.62
4	Stokegate Alycidon	271227	8.5.61
5	Terling Bargain	262167	19.4.60
6	Marlais Hun Adema	292297	21.11.63
7	Holmland Adema	292935	16.1.64
8	Coombevale Jan Dekol	233273	27.7.57
9	Whichford Jantje Uranusz	263417	21.9.60
10	Terling Bonus	253445	8.9.59
11	Buerton Wenword	243727	21.11.58
12	Greatalne Mischevious Boy	264675	2.11.60
13	Ironside Alphonso	274797	6.12.61
14	Marstonmoor Juladema	224819	8.9.55
15	Horwood Jan Squire	275223	28.11.61
16	Scaham Critic	265351	18.1.63
17	Alsopdale Sunbeam 2nd	299487	28.11.64
18	Fintdave Jan Parader	266227	30.11.60
19	Castlerhydd Romeo Tudor	266613	6.11.60
20	Venn Jan	276765	5.2.62
21	Osbournby Triumph	276879	19.12.61
22	Whitegove Marksman RM	257479	12.1.60
23	Hunday Cycher	287549	21.8.63
24	Eynsford Orcades	287921	-
25	Tredene Jan Alidema	289097	28.9.63
26	Lavenham Caprice	279477	25.5.62
27	Holmland Zwarema 17	255923	14.10.59

Table (2.3)

Improved Contemporary Comparisons

Results (from M.M.B. publications)

Sire Number	MILK		FAT %		PROTEIN %		FAT YIELD		PROTEIN YIELD	
	Effective daughters	ICC	Effective daughters	ICC	Effective daughters	ICC	Effective daughters	ICC	Effective daughters	ICC
1	4179.0	+134	4148.2	+0.09	2835.4	-0.02	4148.2	+9.23	2835.4	+3.4
2	3629.3	+182	3578.6	+0.05	2693.7	-0.05	3578.6	+9.3	2693.7	+3.3
3	1229.8	-115	1226.3	+0.02	914.9	+0.00	1226.3	-3.3	914.9	-4.6
4	1420.2	+82	1413.0	+0.22	959.8	+0.13	1413.0	+12.9	959.8	+9.1
5	1552.9	+320	1525.9	-0.10	909.3	-0.12	1525.9	+7.6	909.3	+5.3
6	2748.5	+307	2723.6	-0.05	2146.7	-0.09	2723.6	+10.0	2146.7	+5.3
7	5379.9	+283	5367.8	+0.04	4500.0	-0.02	5367.8	+13.1	4500.0	+8.3
8	2478.7	+365	2467.8	-0.08	1414.8	-0.16	2467.8	+10.3	1414.8	+4.2
9	1288.0	+16	1259.5	-0.04	640.9	-0.03	1259.5	-0.6	640.9	-0.4
10	4566.3	+258	4533.8	-0.04	3699.9	-0.05	4533.8	+7.7	3699.9	+5.8
11	2096.8	+261	2078.8	-0.10	1197.8	+0.01	2078.8	+5.5	1197.8	+8.5
12	1389.9	+278	1365.9	-0.09	756.2	-0.05	1365.9	+6.8	756.2	+6.4
13	2830.5	+211	2818.5	+0.01	2099.6	+0.01	2818.5	+8.6	2099.6	+6.8
14	3814.3	+278	3397.3	-0.07	1815.1	-0.05	3397.3	+7.4	1815.1	+6.5
15	873.6	+124	844.4	+0.01	504.8	+0.04	844.4	+5.1	504.8	+4.5
16	1689.4	+326	1652.3	-0.09	1227.6	-0.05	1652.3	+7.9	1227.6	+8.7
17	4808.9	+395	4794.1	-0.02	4223.7	-0.06	4794.1	+14.3	4223.7	+9.9
18	2491.5	+190	2450.1	+0.01	1457.6	-0.02	2450.1	+7.7	1457.6	+5.4
19	1641.3	+346	1620.0	-0.10	1021.9	-0.04	1620.0	+8.9	1021.9	+8.9
20	1662.9	+105	1637.6	+0.03	1217.5	-0.09	1637.6	+5.7	1217.5	-0.7
21	1926.2	+337	1906.0	+0.02	1269.3	-0.03	1906.0	+13.8	1269.3	+9.5
22	3872.0	+386	3822.0	+0.00	2199.6	-0.03	3822.0	+14.6	2199.6	+10.4
23	1048.0	-135	1048.0	-0.09	751.9	+0.08	1048.0	-9.2	751.9	-0.6
24	1181.2	+220	1166.9	-0.04	835.3	-0.07	1166.9	+6.7	835.3	+3.6
25	2922.8	+338	2922.0	-0.10	2021.0	-0.12	2922.0	+8.1	2021.0	+4.9
26	1706.1	-8	1697.0	+0.10	1211.6	+0.03	1697.0	+4.1	1211.6	+0.6
27	1351.2	+108	1335.7	+0.01	744.2	-0.08	1335.7	+4.4	744.2	-0.2

1. Records made by daughters of any of the set of twenty seven sires that are subject to analysis.

2. Records that were made in the particular herd-year that is to be classified. An average of more than 2.5 independent herd-years were available for classifying any particular herd-year. Three different classes of herd-years were created; low, medium and high. The limits were chosen in such a way as to allow an even distribution of each sire's daughters over levels. A small percentage of herds changed classes from one year to the other. Table (2.4) shows the limits of each classification, the mean and frequency of each class.

b) Herd Size:

Herds were also classified according to numerical size into three different size classes containing approximately equal numbers of herds. The average numbers of heifers recorded over four years was taken as the indicator of herd size. Table (2.5) shows the limits, class means and frequencies for each size class.

c) Calving Season of Herds:

This classification was made in such a way as to exclude calvings made in July, August and January because calvings made in these months are not good indicators of the usual pattern of calving. Herds with a majority of heifers calving in February to June inclusive were considered to be spring calving herds.

Those with a majority of heifers calving in September to December inclusive were deemed to be winter calving herds. Herds with equal numbers of heifers calving in both periods were included in the spring calving group. Table (2.6) shows the means and frequencies of each class.

d) Age at calving groups:

For the purpose of looking at sire X calving age interactions, calving ages were classified into three groups as follows:

Group 1 : 22 to 28 months of age inclusive

Group 2 : 29 to 32 " " "

Group 3 : 33 to 38 " " "

Table (2.7) shows the distribution of heifers over calving ages and seasons of calving.

Table (2.4)

Limits, Means and Frequencies of Level of Herd Production Classes

Level	Limit (Kg)	Mean (Kg)	No. of Animals
low	less than 4120	3886	13198
medium	4120-4635	4404	15855
high	more than 4635	5016	13648

Table (2.5)

Limits, Means and frequencies of Herd Size Classes

Level	Limits	Class Mean (Kg)	Frequency
Small	less than 15	4383	19115
Medium	15-25	4453	11940
Large	More than 25	4553	11646

Table (2.6)

Means and Frequencies of Herd Season of Calving Class

Class	Mean (Kg)	Frequency
Spring	4242	9836
Winter (low)	3859	8843
Winter (medium)	4427	12496
Winter (high)	5066	11526

Table (2.7)

The Distribution of Heifers in Different Age Groups and Seasons

	Calving Age Group			
Season	1	2	3	Total
1	3630	3428	2095	11153
2	2083	3051	1810	6944
3	7500	6712	10392	24604
Total	15213	13191	14297	42701

Section 3

Statistical Methods

Because of the large amount and unbalanced nature of the data it was found necessary to use a least squares programme to obtain unbiased estimates of constants and to calculate the appropriate sums of squares for tests of significance. The following model (2.1) was chosen from the least-squares programme of Harvey (1972) with the following objectives in mind:

1. To provide an estimate of sire X herd-year-season interactions. The problem with this class of interactions is that the number of possible sire X herd-year-season subclasses is very large and only a small percentage of them is filled.
2. To test for differences in the age-yield relationship between different sire families which can be taken as an indication of differences in the rate of maturity between sire groups. This was done by fitting an overall partial regression and a within sire regression on age at calving in months which should show if there are differences between sires over and above the overall regression.

The specific model used is the following:

$$Y_{ijkl} = \mu + H_i + S_j + T_k + (HS)_{ij} + b_1(X_1-\bar{X}_1) + b_2(X_2-\bar{X}_2) + b_{2j}(X_2-\bar{X}_2) + b_3(X_2-\bar{X}_2)^2 + b_{3j}(X_2-\bar{X}_2)^2 + E_{ijkl} \quad (2.1)$$

Where:

Y_{ijkl} = The observation made on the lth heifer in the kth month by the jth sire and in the ith herd-year-season

μ = Overall mean.

H_i = The random effect of the ith herd-year-season.

S_j = The fixed effect of the jth sire, j = 1, 27

T_k = The fixed effect of the kth month within seasons.

$(HS)_{ij}$ = The interaction between the ith herd-year-season and the jth sire.

b_1 = The linear overall regression coefficient on length of lactation (X_1).

b_2, b_{2j} = The linear overall regression coefficient on age at calving (X_2) and the linear regression coefficient within the jth sire respectively.

b_3, b_{3j} = The quadratic overall regression coefficient on age at calving and the quadratic regression coefficient within the jth sire.

E_{ijkl} = The random error term.

A version of model (2.1) that does not include within sire regressions on age at calving was used to obtain estimates of sire effects within different environments i.e. within levels of herd production, within herd calving seasons (spring or winter) and within herd size class.

Length of lactation was corrected for using a linear overall partial regression on the number of days which varied between 285 and 305 days. Seasons were taken out with herd-years and months were corrected for within seasons. The effect of sires was taken to be fixed because of the fact that all of them are old selected sires that have been brought back into service after the initial process of selection. No particular pressure was applied in the choice of herds to be included in the analysis apart from the condition that each herd must have daughters by more than one of the twenty seven sires.

A secondleast-squares mixed model was used in order to look at the interactions of sires with three different classifications of herd environment. These were as follows:

1. Sire X level of herd production (Three levels of production: low, medium and high).
2. Sire X season of herd calving (spring and winter calving herds)
3. Sire X size of herd (Three herd size classes: small, medium and large).

The model was also used to fit within herd class regressions on age at calving to look at the change in age-yield relationship with increasing level of herd production, numerical size and with changes from spring to winter calving herds. Harvey's (1972) programme was used to complete the analysis. The model was:

$$Y_{ijklm} = \mu + V_c + H_{ij} + S_k + T_i + (VS)_{ik} + (HS)_{ijk} + b_1(X_1-\bar{X}_1) + b_2(X_2-\bar{X}_2) + b_{2i}(X_2-\bar{X}_2) + b_3(X_2-\bar{X}_2)^2 + b_{3i}(X_2-\bar{X}_2)^2 + E_{ijklm} \quad (2.2)$$

Y_{ijklm} = The observation made on the mth heifer calving in the lth month by the kth sire in the jth herd-year-season and the ith class of herd.

μ = Overall mean.

V_i = The fixed effect of the ith class of herd.

H_{ij} = The random effect of the jth herd-year-season within the ith class of herd.

S_k = The fixed effect of the kth sire K = 1, 27

T_l = The fixed effect of the lth month within seasons

$(VS)_{ik}$ = The interaction between the ith class of herd with the kth sire.

$(HS)_{ijk}$ = The interaction between the jth herd-year-season within the ith class of herd and the kth sire.

b_1 = The overall partial linear regression coefficient on length of lactation (X_1).

b_2, b_{2i} = The overall partial linear regression coefficient on age at calving (X_2) and the within the ith class of herd linear regression respectively.

b_3, b_{3i} = The overall partial quadratic regression coefficient on age at calving (X_2) and the within the ith class of herd quadratic regression respectively.

E_{ijklm} = The random error term.

A third least-squares mixed model was used primarily to estimate the interaction between age at calving groups and seasons of calving. Other estimates obtained from the model were the herd-year-sire X age at calving and the herd-year-sire X season of calving interactions. The model was designed to correct for herd-year-sire effects and fit seasons, age groups, and all two way interactions between these effects. In order to accommodate this model into Harvey's (1972) programme it was found necessary to run it in a number of stages. These were as follows:

$$Y_{jkl} = \mu + S_j + A_k + (SA)_{jk} + E_{jkl} \qquad (2.3)$$

where:

Y_{ijkl} = The observation made on the lth heifer calving in the jth season and in the kth age group.

μ = Overall mean

S_j = The fixed effect of the jth season j = 1, 3.

A_k = The fixed effect of the kth age group k = 1, 3

$(SA)_{ij}$ = The interaction of the jth season of calving with the kth age group.

E_{jkl} = The random error term.

This run ignores the random effect of herd-year-sires and fits all other fixed effects. The residual sum of squares (E_1) from this run is equivalent to: $Y^1Y - R(\mu,S,A,SA)=E_1$, where Y^1 and Y are a row and a column vector of observations, and the symbol (R) stands for the reduction in sums of squares.

$$Y_{ijkl} = \mu + H_i + S_j + A_k + (SA)_{jk} + E_{ijkl} \qquad (2.4)$$

where Y_{ijkl} = The observation made on the lth heifer calving in the kth age group and the jth season and in the ith herd-year-sire group.

H_i = The random effect of the ith herd-year-sire.

The rest of the notation is above. In this run the $\mu + H_i$ equations were absorbed while all interactions involving herd-year-sires are ignored. The residual sum of squares (E_2) from this run is:

$Y^1Y\ R(\mu,H,S,A,SA) = E_2$

$$Y_{ijkl} = \mu + T_i + A_j + (SA)_{jk} + E_{ijkl} \qquad (2.5)$$

where T_i is the ith herd-year-sire-season combination. This effect was absorbed in this run. The residual sum of squares (E_3) is $Y^1Y - R(\mu,H,S,A,HS,SA) = E_3$

$$Y_{ijkl} = \mu + B_i + S_j + (SA)_{jk} + E_{ijkl} \qquad (2.6)$$

Where B_i is the ith herd-year-sire-age group combination. This effect was absorbed in this run. The residual sum of squares (E_4) is $Y^1Y - R(\mu, H, S,A,HA, SA) - E_4$

$$Y_{ij} = \mu + V_i + E_{ij} \qquad (2.7)$$

where V_i is the ith herd-year-sire-season-age group. The effect of herd-year-sire-season-age group (V_i) was absorbed in this run and this resulted in the removal of the three way interaction (herd-year-sire X season X age) which is not actually included in the model . The effect of this is that the error term (E_5) is likely to be slightly under estimated. The residual sum of squares (E_5) from this run is:

$Y^1Y - R(\mu, H, S, A, HS, HA, HSA) = E_5$

From the first two runs the sum of squares of herd-year-sires can be calculated as $E_1 - E_2$. The effect of herd-year-sire X season interaction is estimated from $E_2 - E_3$, while that of herd-year-sire X age interaction is obtained from $E_2 - E_4$ and the error sum of squares can be taken to be approximately equivalent to E_5. The overall model is as follows:

$$Y_{ijkl} = \mu + H_i + S_j + A_k + (HS)_{ij} + (HA)_{ik} + (SA)_{jk} + E_{ijkl} \qquad (2.8)$$

where:

$(HS)_{ij}$ = The interaction of the ith herd-year-sire with the jth season.

$(HA)_{ik}$ = The interaction of the ith herd-year-sire with the kth age group.

The effect of months within seasons and within age groups and the effect of length of lactation were corrected for using partial regressions.

Two models were used to analyse estimates of least-squares means of sires obtained from different classifications of data. The first model was used to obtain estimates of both the sire X season of herd calving and sire X level of production interactions. The model used is the following:

$$Y_{ijkl} = \mu + T_i + L_{ij} + S_k + (TS)_{ik} + (SL)_{kij} + E_{ijkl} \qquad (2.9)$$

where:

Y_{ijkl} = The least-squares mean of the kth sire in the jth level of production within the ith herd calving season.

μ = Overall mean.

T_i = The fixed effect of the ith herd calving season, i = 1, 2.

L_{ij} = The fixed effect of the jth level of production, j = 1, 3.within the ith herd calving season.

S_k = The fixed effect of the kth sire, k = 1, 27

$(TS)_{ik}$ = The interaction between the ith herd calving season and the kth sire.

$(SL)_{kij}$ = The interaction between the kth sire and the jth level of production within the ith herd calving season.

E_{ijkl} = The random error term.

It was not possible to estimate the error term from the observations directly due to the lack of sufficient numbers of degrees of freedom. The error mean square was then estimated using the following procedure: The expected mean square for error is $\frac{E_2}{n}$ where E_2 is the error mean square obtained by applying model (2.2) and $n = \frac{S \times L}{\sum \frac{1}{h}}$

where: S = The number of sires (27)

L = The number of levels (4)

h = The effective number for each sire in particular levels.

The second model used to analyse estimates of sires' least-squares means is a reduced form of model (2.9) above. It was used on estimates obtained from data divided into three classes of herds only. The specific model is:

$$Y_{ijk} = \mu + S_i + L_j + (SL)_{ij} + E_{ijk} \qquad (2.10)$$

where Y_{ijk} = The estimate of least-squares means of the ith sire.

μ = Overall mean.

S_i = The fixed effect of the ith sire.

L_j = The fixed effect of the jth class of herd.

E_{ijk} = The random error term.

A similar method to that explained in connection with model (2.9) was used to estimate the error mean square.

Where linear and quadratic average and individual class regressions are fitted, prediction equations used in plotting graphs are described by Harvey (1972).

Section 4: RESULTS

In this section the results of the analyses on the twenty seven wide-spread use sires will be presented. The main areas covered by the results are as follows:-

1. Sire X Herd-year-season interaction: The model used to estimate this interaction also includes within sire regressions on age at calving.
2. Sire X level of herd production: This is a study based on classifying the herd-years according to three criteria. These are:

a. Herd-year level of milk production
b. The calving season of herds.
c. The size of herds

3. Interactions involving age at calving: These are studied under four main headings:

a. Sire X age interactions
b. Age X level of herd production interaction
c. Age X season interaction
d. Two and three way interactions involving herds, years, sires and age.

1. Sire X herd-year-season interactions:

In order to estimate the sire X herd-year-season mean squares for milk yield, fat percent, and fat yield, model (2.1) was run on a set of more than 42000 records. The procedure was repeated on a set of more than 27000 records to estimate the corresponding mean squares for protein per cent and protein yield. As will be shown later yield traits tend to exhibit large changes in the within and between sire variance from one level to the other. In order to achieve greater homogeneity of variance over different levels the

scale of the analysis was changed by transforming the data on yield traits to natural logarithms.

The results of this analysis are shown in tables (2.8 - 2.10). Table (2.8) gives the mean squares and F-ratio values for milk yield, fat per cent and fat yield using an additive model i.e. without transforming the data. It is clear that out of a possible number of more than 300000 sire X herd-year-season sub-classes only 7610 are filled. The F-ratios for the sire X herd-year-season interaction for milk yield and fat yield are below significance level at 5%. The F-ratio for fat per cent is just above significance level which, for a test where both the numerator and denominator have relatively large numbers of degrees of freedom, is close to unity. However for all three traits the actual magnitude of the interaction mean square is small and even in the case of fat per cent, where it is significant, it is not likely to account for more than 1-2% of the total variance in milk yield.

Table (2.9) shows milk yield and fat yield mean squares and F-ratios resulting from applying model (2.1) after transforming the data to logarithms. With the greater homogeneity of variance introduced by the transformation of the data, the sire X herd-year-season interaction F-ratio fell further from 0.91 to 0.85 in the case of milk yield and that of fat yield from 0.89 to 0.83. However the purpose of the transformation was mainly to study the effect of a multiplicative model on age at calving regressions. This will be taken up later when the results of age interactions are presented. Table (2.10) gives the results of applying model (2.1) to the protein set of data. The mean squares

TABLE (2.8)

Analysis of Variance Table: Sire X Herd-year-season Interaction and Regressions Within Sires on Age at Calving Using Model (2.1) †

Source	D.F	Milk Yield M.S.	Milk Yield F	Fat Per Cent M.S.	Fat Per Cent F	Fat Yield M.S.	Fat Yield F
Herd-year-seasons	11968	1499288	4.214**	0.2151	2.805**	2486.58	5.141**
Sires	26	5773701	16.228**	1.9908	24.241**	7760.43	16.043**
Sires X Herd-year-seasons	7610	321876	0.905	0.0821	1.071**	428.34	0.886
Months in season 1	3	2090709	5.876**	0.0510	0.666	2427.91	5.019**
Months in season 2	3	730759	2.054	0.2745	3.580*	1267.25	2.620
Months in season 3	3	5695074	16.007**	0.0586	0.765	9535.36	19.713**
Length of lactation (linear)	1	1144685206	3217.368**	0.3860	5.033	1512486.65	3126.787**
Calving age (linear)	1	156036907	438.573**	0.5419	7.067	250170.83	517.182**
Calving age within sires (linear)	26	574719	1.615*	0.1194	1.557*	671.95	1.389
Calving age (quadratic)	1	3251574	9.139**	0.1795	2.340	2490.32	5.148*
Calving age within sires (quadratic)	26	290066	0.815	0.0495	0.645	477.72	0.988
Remainder	23032	355783		0.0767		483.72	

† The effect of sires is tested against the sire X herd-year-season mean square. The rest of the terms are all tested against the remainder.

* P<0.05 ** P<0.01

TABLE (2.9)

Analysis of Variance Table: Sire X Herd-year-season interaction and Regressions Within Sires on Age at Calving (logarithmic scale) †

		Log (milk yield)		Log (fat yield)	
Source	D.F	M.S	F	M.S	F
Herd-year-seasons	11968	0.0149	4.275**	0.0176	5.167**
Sires	26	0.0849	24.257**	0.1496	52.757**
Sires X Herd-year -season	7610	0.0030	0.849	0.0028	0.833
Months in season 1	3	0.0199	5.696**	0.0168	4.948**
Months in season 2	3	0.0075	2.150	0.0082	2.403
Months in season 3	3	0.0761	21.777**	0.0891	26.188**
Length of lactation (linear	1	15.2671	4367.822**	14.7478	4332.388**
Calving age (linear	1	1.5680	448.584**	1.7707	520.164**
Calving age within sires (linear)	26	0.0059	1.692*	0.0049	1.445
Calving age (quad-ratic)	1	0.0504	14.410**	0.0305	8.960**
Calving age within sires (quadratic)	26	0.0028	0.810	0.0035	1.017
Remainder	23032	0.0035	-	0.0034	-

† The sire effect is tested against the sire X herd-year-season mean square. All other terms are tested against the remainder.

TABLE (2.10)

Analysis of Variance Table: Sires X Herd-year-season and Regressions Within Sires on Age at Calving (additive and logarithmic scales) †

Source	D.F	Protein per cent M.S	Protein per cent F	Protein yield (kg) M.S	Protein yield (kg) F	Log (Protein yield) M.S	Log (Protein yield) F
Herd-year-seasons	7764	0.0600	2.370**	2094.83	6.723**	0.0215	7.237**
Sires	26	0.6466	24.266**	20501.37	65.800**	1.4921	231.243**
Sires X Herd-year-seasons	4991	0.0266	1.053**	438.55	1.408**	0.0065	2.170**
Months in season 1	3	0.1934	7.645**	612.00	1.964	0.0067	2.255
Months in season 2	3	0.3378	13.350**	1767.42	5.673**	0.0179	6.004**
Months in season 3	3	0.5116	20.220**	1538.61	4.938**	0.0220	7.390**
Length of lactation	1	0.2799	11.026**	676639.01	2171.685**	8.7779	2951.939**
Calving age (linear)	1	0.1388	5.485*	104462.03	335.273**	1.0090	339.325**
Calving age within sires (linear)	26	0.0247	0.976	407.26	1.307	0.0042	1.411
Calving age (quadratic)	1	0.0605	2.389	441.73	1.418	0.0118	3.969*
Calving age within sires (quadratic)	26	0.0200	0.792	306.05	0.982	0.0036	1.201
Remainder	14729	0.0253	-	311.57	-	0.0030	-

† The sire effect is tested against the sire X herd-year-season mean square. All other terms are tested against the remainder.

and F-ratios of protein yield after the data were transformed to logarithms are also given. On the additive scale both protein per cent and protein yield exhibit significant sire X herd-year-season interactions. The F-ratio of the protein yield interaction mean square is the highest for all traits (1.41). After transformation the protein yield F-ratio rises significantly to 2.17. Although significant, the interaction mean square is far less important than correcting for length of lactation or age at calving or the other major components of the model. However, as will be shown later, protein yield seems to exhibit higher interactions than other traits.

As with the milk and fat set of data, the protein set is also highly non orthogonal. The numbers of full sire X herd-year-season subclass does not exceed 2-3% of the total possible number. It is clear that a broader classification of environments is necessary to improve the reliability of the results. Analyses using such classifications will be presented in the following sections.

In this and the following chapters all estimates of yield traits will be given in kilogrammes. Significance levels will be indicated as follows unless otherwise stated.

* Significant at 5%

** Significant at 1%

2. Sire X Level of Herd Interactions

Under this heading three different sets of analyses will be presented. These are:

a. Sire X level of Herd Production: The level of herd production here is defined by the herd-year average milk production as computed from herd years and progeny groups other than those included in the analyses. Three herd levels of production were created averaging 3886, 4404 and 5016 kilogrammes of milk. The results on sires X level of production will be presented in two ways, firstly as a combined cross classification analysis of variance using model (2.2). This model accounts for both sire X herd-year season interaction and sire X level of herd production interaction. Table (2.11) shows the analysis of variance on milk yield, fat per cent and fat yield. The sire X level of herd production is significant at the 5% level for both milk yield fat yield. However, in both cases the interaction mean square is very small compared with other terms in the model. In both traits the sire X herd-year-season term is also just above significance level but does not account for much of the variation in either trait. The sire X level of herd production interaction mean square is not significant at 5% although that of sire X herd-year-season is just above significance level. Table (2.12) shows the results from applying the same model on milk yield and fat yield data transformed to logarithms. The F-ratio for the sire X level of herd production for both traits drops slightly after transformation indicating that some of the interaction is accounted for by the change of scale. In both traits the sire X herd-year-season interaction remains significant but relatively small.

Table (2.13) shows the results of the analysis using model (2.2) on the protein set. A similar trend to that noted

TABLE (2.11)

Analysis of Variance Table: Sires X Level of Herd Production Interactions and Regressions within levels on Age at Calving †

Source	D.F	Milk Yield M.S.	Milk Yield F	Fat Per Cent M.S	Fat Per Cent F	Fat Yield M.S.	Fat Yield F
Levels	2	1535591125	1832.202**	4.1367	19.364**	2384028.32	1685.756**
Herd-year-seasons within levels	11966	838112	2.354**	0.2136	2.786**	1414.22	2.923**
Sires	26	6314063	15.930**	2.7364	33.337**	8115.34	15.440**
Levels X Sires	52	541065	1.365*	0.0677	0.825	705.17	1.342*
Sires X Herd-year-seasons	7521	396360	1.113**	0.0821	1.071**	525.61	1.086**
Months in season 1	3	2386802	6.705**	0.0562	0.732	2839.63	5.870**
Months in season 2	3	681263	1.914	0.2790	3.639*	1218.51	2.519
Months in season 3	3	5960531	16.744**	0.0557	0.727	9840.31	20.341**
Length of lactation	1	1138420831	3197.906**	0.3764	4.909*	1505561.21	3112.164**
Calving age (L)	1	189239775	531.588**	0.8383	10.933**	308496.46	637.697**
Calving age (l) within levels	2	158957	0.447	0.0977	1.274	382.05	0.790**
Calving age (Q)	1	2546525	7.153**	0.2546	3.321	1653.14	3.417
Calving age (Q) within levels	2	3639	0.010	0.0602	0.786	57.41	0.119
Remainder	23117	355990	-	0.0767	-	483.77	-

† The effect of levels is tested against the herd-year-seasons within levels mean square. Both sires and sires X level of herd are tested against the sire X herd-year-season interaction. All other terms are tested against the remainder.

TABLE (2.12)

Analysis of Variance Table: Sire X Herd Production and Within Levels Regressions on Age at Calving (logarithmic scale) †

Source	D.F	Log(Milk Yield) M.S.	F	Log (Fat Yield) M.S.	F
Levels	2	15.1113	1923.217**	16.8956	1681.371**
Herd-year-seasons within levels	11966	0.0079	2.248**	0.0100	2.951**
Sires	26	0.0623	15.871**	0.0554	14.779**
Levels X sires	52	0.0048	1.225	0.0045	1.202
Sires X Herd-year -seasons	7521	0.0039	1.122**	0.0037	1.102**
Months in season 1	3	0.0233	6.656**	0.0201	5.916**
Months in season 2	3	0.0073	2.084	0.0080	2.348
Months in season 3	3	0.0786	22.481**	0.0913	26.823**
Length of lactation	1	15.1863	4344.240**	14.6759	4312.596**
Calving age (L)	1	1.8613	532.463**	2.1455	630.480**
Calving age (L) within levels	2	0.0115	3.290*	0.0049	1.439
Calving age (Q)	1	0.0409	11.708**	0.0215	6.329**
Calving age (Q) within levels	2	0.0001	0.028	0.0012	0.347
Remainder	23117	0.0035	-	0.0034	-

† The effect of levels is tested against the herd-year-seasons within levels mean square. Both sires and sires X level of herd are tested against the sire X herd-year-season interaction. All other terms are tested against the remainder.

TABLE (2.13)

Analysis of Variance Table: Sires X Level of Herd Production and Regressions Within Levels on Age at Calving (additive and logarithmic scales) †

Source	D.F	Protein % M.S	Protein % F	Protein Yield M.S.	Protein Yield F	Log (Protein Yield)	Log (Protein Yield)
Levels	2	1.7974	30.862**	1120649.33	1318.439**	10.7596	1311.730**
Herd-year-seasons	7762	0.0582	2.302**	849.98	2.728**	0.0082	2.759**
Sires	26	1.0333	38.340**	3675.50	10.822**	0.0358	10.796**
Levels X sires	52	0.0327	1.211	356.55	1.050	0.0038	1.153
Sires X Herd-year-seasons	4908	0.0270	1.066**	339.63	1.090**	0.0033	1.115**
Months in season 1	3	0.1651	6.525**	583.85	1.874	0.0068	2.303
Months in season 2	3	0.3352	13.252**	1728.28	5.547**	0.0174	5.857**
Months in season 3	3	0.5405	21.369**	1588.70	5.099**	0.0223	7.502**
Length of lactation	1	0.2914	11.519**	671818.48	2156.269**	8.7011	2926.739**
Calving age (L)	1	0.2454	9.700**	133423.43	428.236**	1.2717	427.756**
Calving age (L) within levels	2	0.0399	1.576	172.72	0.554	0.0064	2.157
Calving age (Q)	1	0.0424	1.675	853.74	2.740	0.0172	5.769*
Calving age (Q) within levels	2	0.0071	0.282	37.20	0.119	0.0002	0.051
Remainder	14808	0.0253	-	311.57	-	0.0030	-

† The effect of levels is tested against the herd-year-seasons within levels mean square. Both sires and sires X level of Herd production are tested against the sire X Herd-year-season interaction. All other terms are tested against the remainder.

for milk and fat yield is also apparent here. Both sets of interactions mean squares are small although that of sire X herd-year-season interaction is statistically significant in both protein per cent and protein yield. There is no significant change in this situation which is contrary to the findings noted in table (2.10) above where transformation resulted in a significant increase in the importance of sire X herd-year-season interaction.

b. A separate analysis was carried out on each herd level of production using model (2.1) in order to account for differences in variance between the three levels of production and to provide three independent estimates of sires' least-squares means. These estimates are plotted in figures (2.1 - 2.3). Some of the outlying sires are identified on the scatter diagrams by their serial numbers as given in table (2.2) and their constant estimates and effective numbers of daughters are shown in table (2.15). Table (2.14) shows the product moment correlations between sires' least-squares means in the three levels of herd production. The correlations for milk yield, fat per cent, fat yield and protein per cent are all high and range between 0.77 and 0.94. However the correlations for protein yield are lower and range between 0.64 and 0.80. This is unexpected since the F-ratio for sire X level of herd production shown in table (2.14) is not higher than that obtained for other traits. Table (2.16) shows the degress of freedom, between sire mean squares and within sire mean squares from the analysis using model (2.1) on separate production levels. This table is produced mainly to show the change in variance that accompanies the rise in the mean level of herd production. The error mean square for milk yield rises from

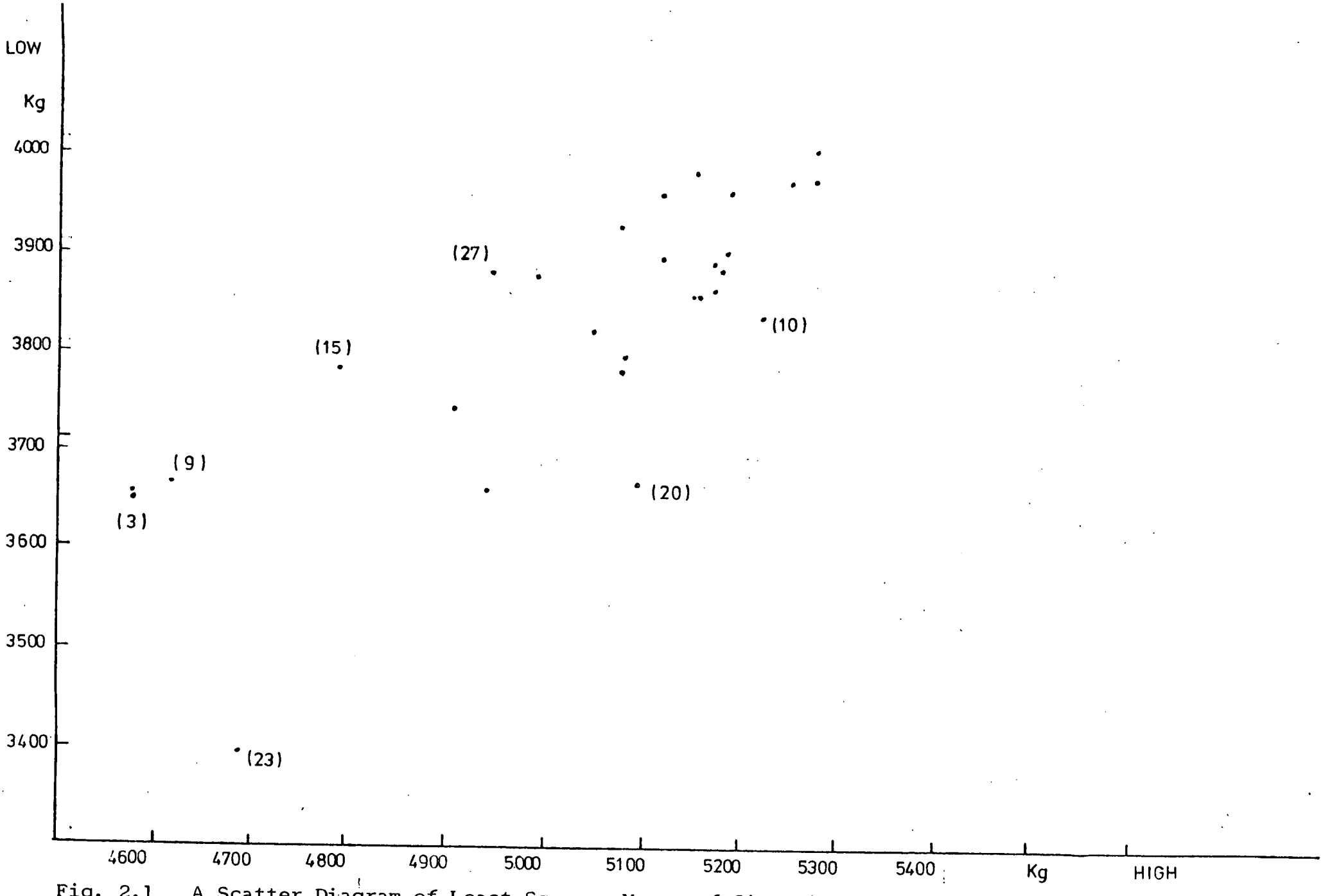


Fig. 2.1 A Scatter Diagram of Least Squares Means of Sires in the Low and High Levels of Production (Milk Yield).

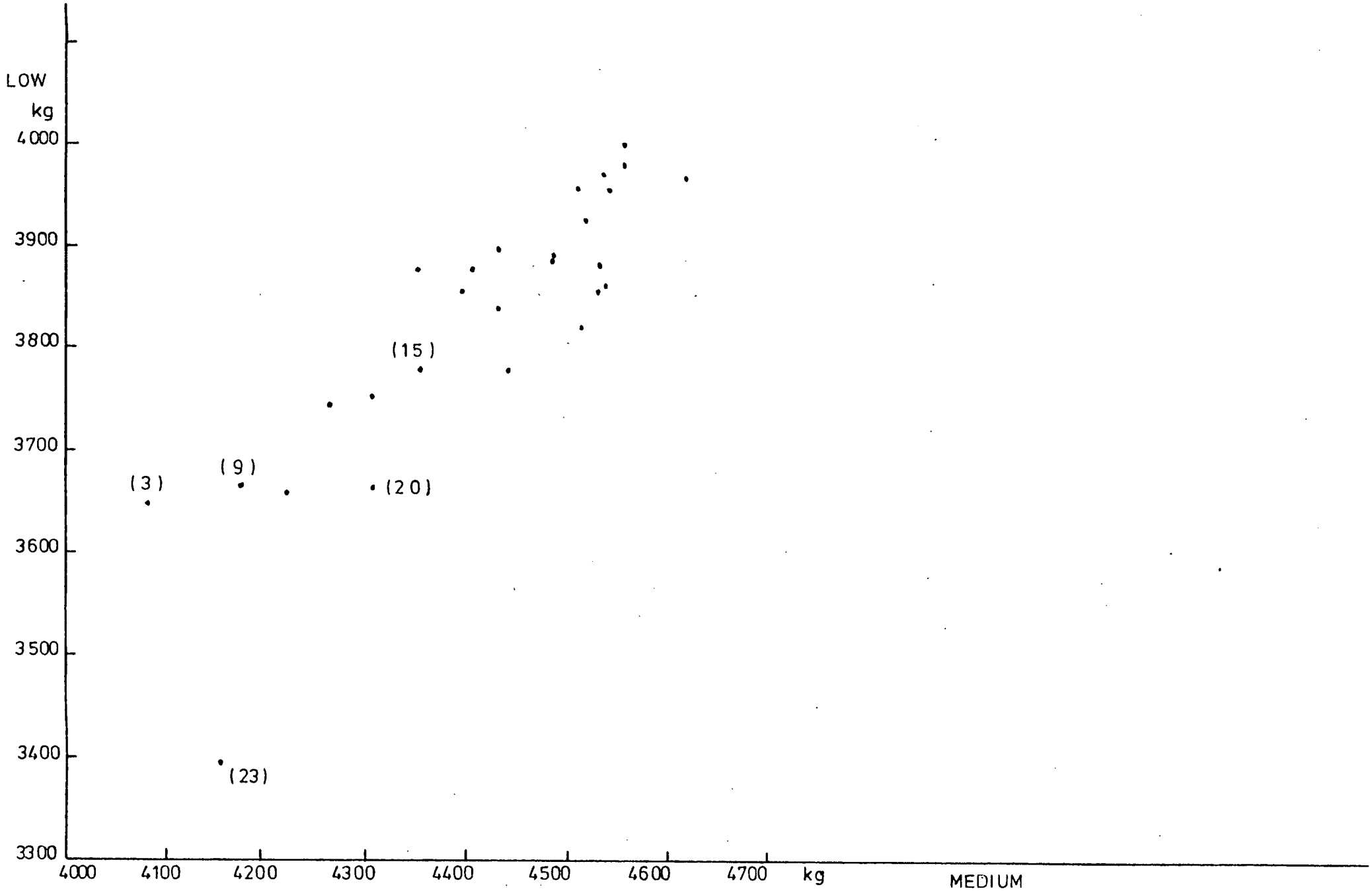


Fig. 2.2 A Scatter Diagram of Least Squares Means of Sires in the Low and Medium Levels of Production (Milk Yield).

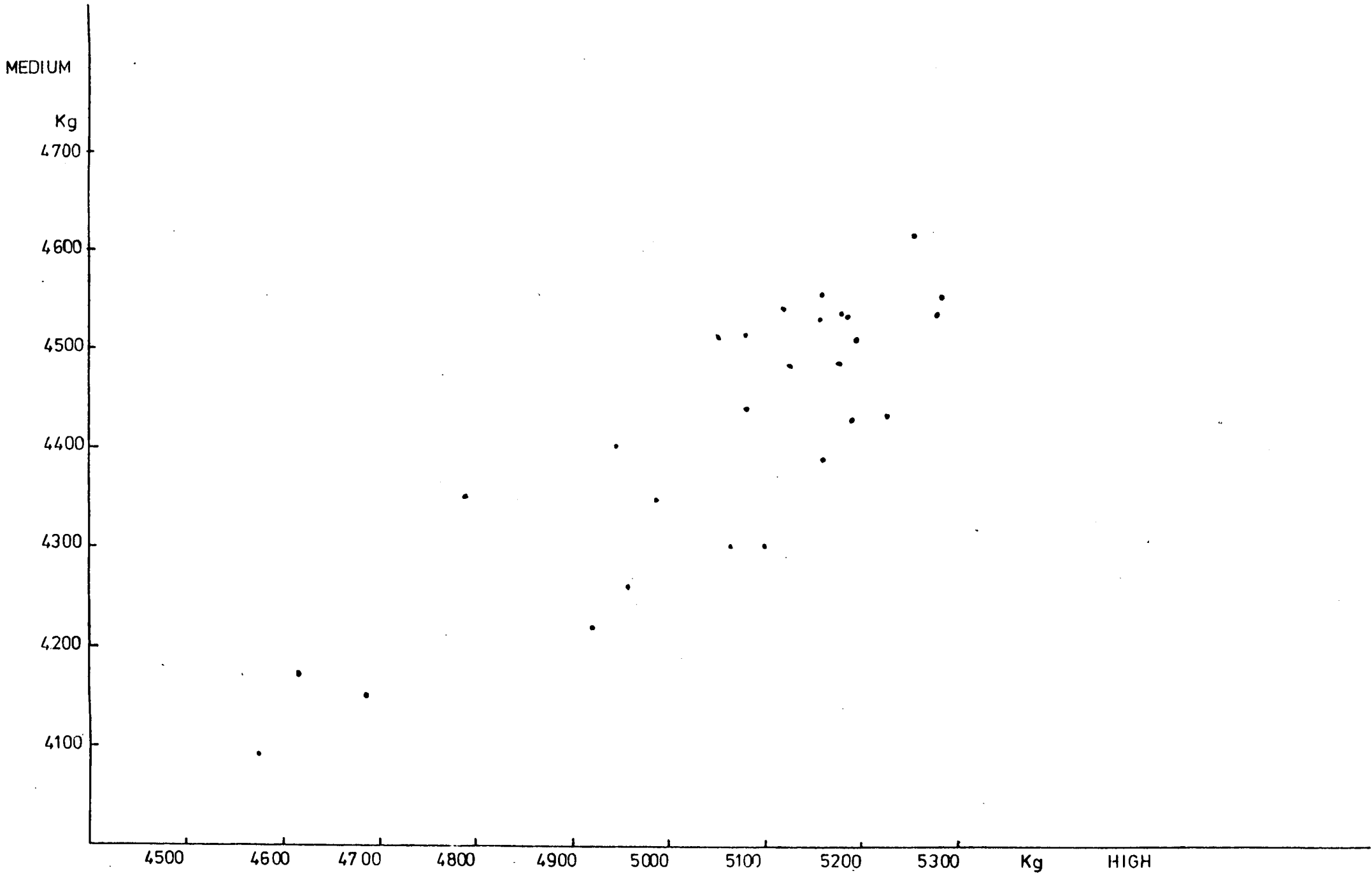


Fig. 2.3 A Scatter Diagram of Least Squares Means of Sires in the Medium and High Levels of Production (Milk Yield).

TABLE (2.14)

Product Moment Correlations Between
Sire Least Square Means in
Different Levels of Production

		medium	high
Milk yield	low	0.8646	0.7667
	medium		0.8565
Fat per cent	low	0.9358	0.8476
	medium		0.9087
Fat yield	low	0.8776	0.8524
	medium		0.8872
Protein per cent	low	0.8674	0.8374
	medium		0.9095
Protein yield	low	0.6402	0.6596
	medium		0.7967

TABLE (2.15)

Effective Numbers of Daughters, Constant Estimates and Standard Errors for Seven Outlying Sires

Serial Number of sire	Level of Production					
	Low		Medium		High	
	Effective Number	Constant Estimate ± Standard error	Effective Number	Constant Estimate ± Standard Error	Effective Number	Constant Estimate ± Standard error
3	9842	-179 ± 100	87	-333 ± 83	38	-480 ± 125
9	162	-163 ± 44	128	-241 ± 57	45	-440 ± 101
10	89	8 ± 61	127	18 ± 59	183	168 ± 53
15	45	-49 ± 92	56	-62 ± 93	13	-440 ± 101
20	146	-164 ± 47	140	-111 ± 55	90	42 ± 75
23	30	-433 ± 102	92	-262 ± 68	134	-368 ± 63
27	70	50 ± 69	51	-11 ± 90	50	-112 ± 98

TABLE (2.16)

Analysis of Variance Table: Estimates of Within and Between Sire Mean Squares Resulting from Applying Model (2.1) on Separate Levels of Herd Production †

	Low				Medium				High			
	Error		Sires		Error		Sires		Error		Sires	
	D.F	M.S.	M.S.	F.	D.F	M.S.	M.S.	F.	D.F	M.S.	M.S.	F.
Milk yield	6902	280194	1921520	6.2	8668	345805	3140158	7.8	7527	436203	3340373	7.0
Fat per cent	6902	0.0752	0.7294	9.4	8668	0.0745	1.1842	14.6	7527	0.0805	1.2028	13.8
Fat yield	6902	374.59	2569,16	6.3	8668	470.73	3510.37	6.6	7527	597.58	5264.38	8.4
Protein per cent	3889	0.0254	0.3126	12.6	5913	0.0251	0.4721	16.6	4986	0.0253	0.4263	15.9
Protein yield	3889	242.66	1239.94	4.6	5913	300.12	1873.28	5.5	4986	377.75	1591.91	4.0

† The degrees of freedom for sires in all three levels are 26.

280000 in the low level to 436000 in the high, a rise of about 55%. There is a similar rise in the between sire mean square from 1920000 (F = 6.20) in the low level to 3340000 (F = 7) in the high level, a rise of about 75%. If we assume equal numbers of effective daughters in the three levels, which is a reasonable assumption, it is clear that the between sire component of variance has increased relatively more than the within sire variance. A similar argument can also be applied to both fat and protein yield which exhibited large changes in variance. However in the case of fat per cent and protein per cent the changes in variance were much smaller.

b. Sire X Calving season of Herd Interactions

In this analysis the sire X level of herd interaction is studied by first classifying herds into two major classes of Spring (February to June) and Winter (September to December) calving herds. The winter calving herds are then further classified according to the herd-year mean milk production into low, medium and high levels of production. The spring calving herds were not further split into different levels because of the small number of records involved. The cross-classification analysis of sires X levels was abandoned here in favour of applying model (2.1) on separate levels because of the large shifts in variance demonstrated above. For each sire four separate estimates of its least-square mean were obtained and the resulting 108 estimates were analysed using model (2.9) to provide a check on the statistical significance of the interaction mean square.

Table (2.17) gives the analysis of variance results obtained from applying model (2.1) on the separate herd classes.

TABLE (2.17)

Analysis of Variance Table: Between and Within Sire Mean Squares obtained by Applying Model (2.1) to Separate Herd Classes

Trait	Low Winter				Medium Winter				High Winter				Spring			
	Error		Sires		Error		Sires		Error		Sires		Error		Sires	
	D.F	M.S.	M.S.	F.	D.F	M.S.	M.S.	F	D.F	M.S.	M.S.	F	D.F	M.S.	M.S.	F
Milk Yield	4887	287633	1567876	4.9	6953	354899	2534801	6.3	6497	446178	2919196	5.9	4756	303048	1711598	5.1
Fat per cent	4887	0.0749	0.4666	6.0	6953	0.0737	0.9748	12.1	6497	0.0810	1.0573	12.0	4756	0.0770	0.7068	8.7
Fat yield	4887	384.37	2075.32	4.9	6953	481.00	2793.73	5.2	6497	599.30	4547.38	7.0	4756	429.88	2380.44	5.3
Protein per cent	2756	0.0246	0.2494	10.3	4854	0.0252	0.3555	12.6	4385	0.0253	0.3823	14.5	2787	0.0260	0.2559	9.1
Protein yield	2756	247.46	1031.98	3.8	4854	302.78	1496.32	4.4	4385	384.30	1508.30	3.7	2787	273.51	1067.93	3.6

A similar trend in variance to that noted in table (2.16) is apparent here. The spring calving class comes second to the winter low class in the magnitude of the within and between sire variance similar to the position it occupies with regard to the mean level of production. Tables (2.18) and (2.19) shows the results of the analysis using model (2.9) on the 108 estimates of sires' least-squares means. For all traits, except protein yield, the F-ratiofor the sire X season and the sire X levels within seasons interactions approach significance at the 5% level but the magnitude of the interaction mean square is small. In the case of protein yield, however, the sire X season interaction appears to be of considerable importance, the sire X levels within seasons being less so. This is further emphasized by the product moment correlations between sires' least square means shown in table (2.20). The correlations for milk yield, fat per cent, fat yield and protein per cent are consistently high and range between 0.70 and 0.90. Protein yield correlations between the three winter levels are close to those obtained in the sire X level of herd production analysis above and range between 0.61 and 0.74. However all estimates of correlation involving the spring class of herds are extremely low and range between 0.28 and 0.52. The effective numbers of daughters for the spring class are generally low and two sires which exhibited a large change in value have very small numbers of effective daughters. In figure (2.4) the least-squares means of sires obtained from the low winter and the spring calving classes of herds are plotted against each other. It is clear from the figure that most of the interaction is not due to sires exhibiting large deviations in their mean level of production from one level

TABLE (2.18)

Analysis of Variance Table: Analysing Estimates of Sires' Least-Squares Means According to Model (2.9) (Set 1)

	Milk Yield			Fat Per Cent			Fat Yield		
Source	D.F	M.S	F	D.F	M.S.	F	D.F.	M.S	F
Reduction due to the mean	1	1600.00	0.343	1	20.8291	34715.167**	1	18172.84	4275.962**
Season	1	877344.44	187.982**	1	0.0135	22.5**	1	909.83	214.078**
Sires	26	60910.30	13.051**	26	0.0178	29.667**	26	83.03	19.536**
Levels within seasons	2	9834922.93	2107.256**	2	0.0151	25.167**	2	15132.70	3560.635**
Sires X Seasons	26	6213.28	1.331	26	0.008	1.333	26	7.91	1.861**
Sires X Levels within seasons	52	6093.75	1.306	52	0.0007	1.167	52	6.17	1.452
Remainder	-	4667.17		-	0.0006		-	4.25	

TABLE (2.19)

Analysis of Variance Table: Analysing Estimates of Sires' Least-Squares Means According to Model (2.9) (Set 2)

	Protein %			Protein Yeild		
Source	D.F	M.S		D.F.	M.S.	
Reduction due to the mean	1	0.0009	1.5	1	319449.16	42255.180**
Season	1	0.0044	7.333**	1	978.13	129.382**
Sires	26	0.0140	23.333**	26	39.53	5.229
Levels within seasons	2	0.0115	19.167**	2	11130.36	1472.270**
Sires X Seasons	26	0.0013	2.167**	26	14.26	1.886**
Sires X Levels within seasons	52	0.0007	1.167	52	7.66	1.013
Remainder	-	0.0006		-	7.56	

TABLE (2.20)

Product Moment Correlations Between Estimates of Sires Least-Squares Means in Different Herd Classes

	Medium	High	Spring
Milk yield			
Low	0.8298	0.7559	0.7705
Medium		0.8125	0.8199
High			0.7035
Fat per cent			
Low	0.9059	0.8806	0.9006
Medium		0.8890	0.8992
High			0.8602
Fat yield			
Low	0.7948	0.8293	0.7196
Medium		0.8936	0.8762
High			0.8300
Protein percent			
Low	0.9729	0.8374	0.8493
Medium		0.8874	0.8863
High			0.9413
Protein yield			
Low	0.6128	0.6084	0.2842
Medium		0.7356	0.4255
High			0.5243

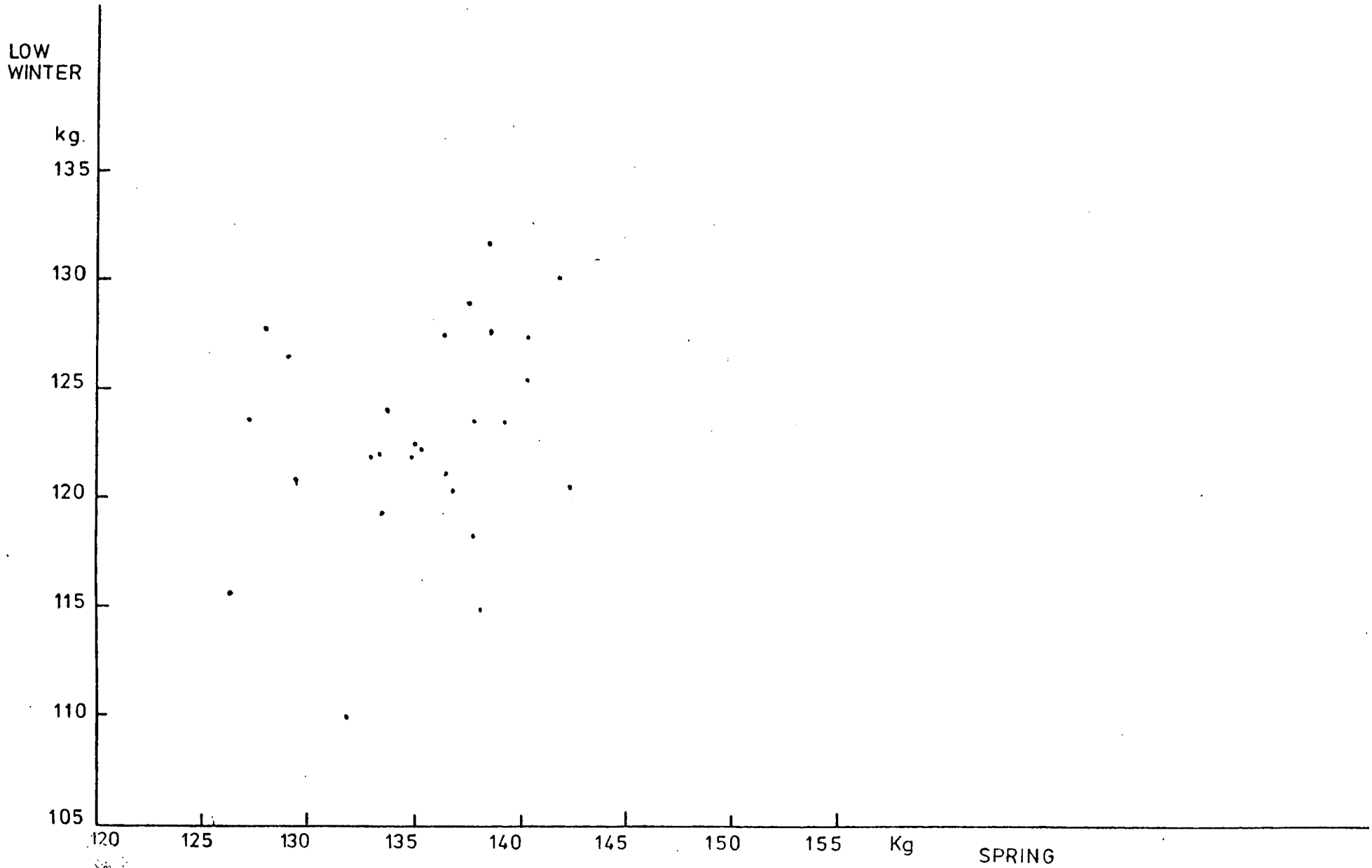


Fig. 2.4 A Scatter Diagram of Sires' Least Squares Means in Spring Herds and Low Winter Herds (Protein Yield).

to the other but rather to a large degree of change in the ranking of sires from one class of herds to the other. Figures (2.5 - 2.7) shows the relationship between least-squares means estimates for milk yield over different herd classes. Unlike figure (2.4) these tend to show a greater range in both levels and a few individual sires showing large changes in their deviations from the mean of the herd classes.

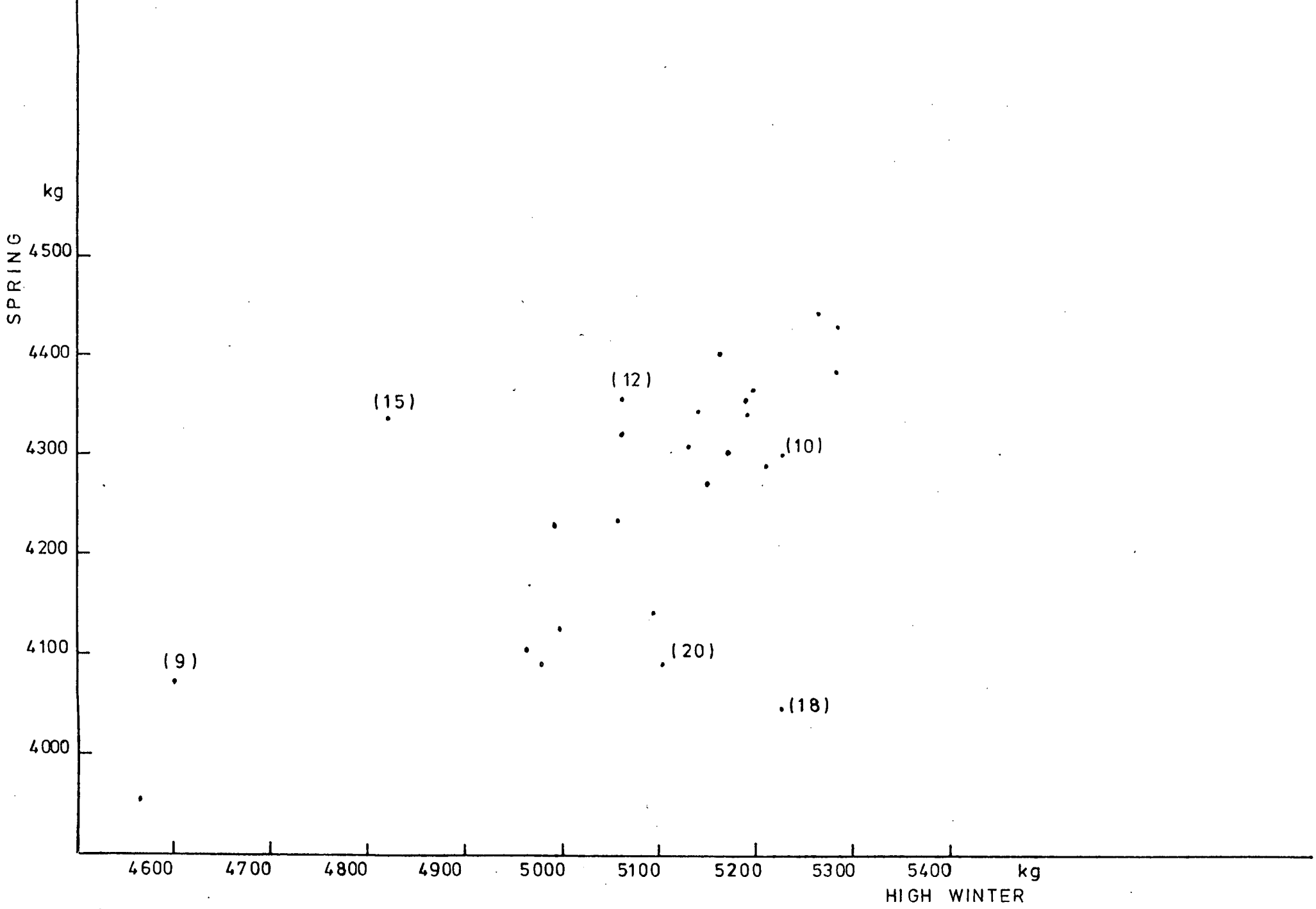


Fig. 2.5 A Scatter Diagram of Least Squares Means of Sires in the Spring and High Winter Herds (Milk Yield).

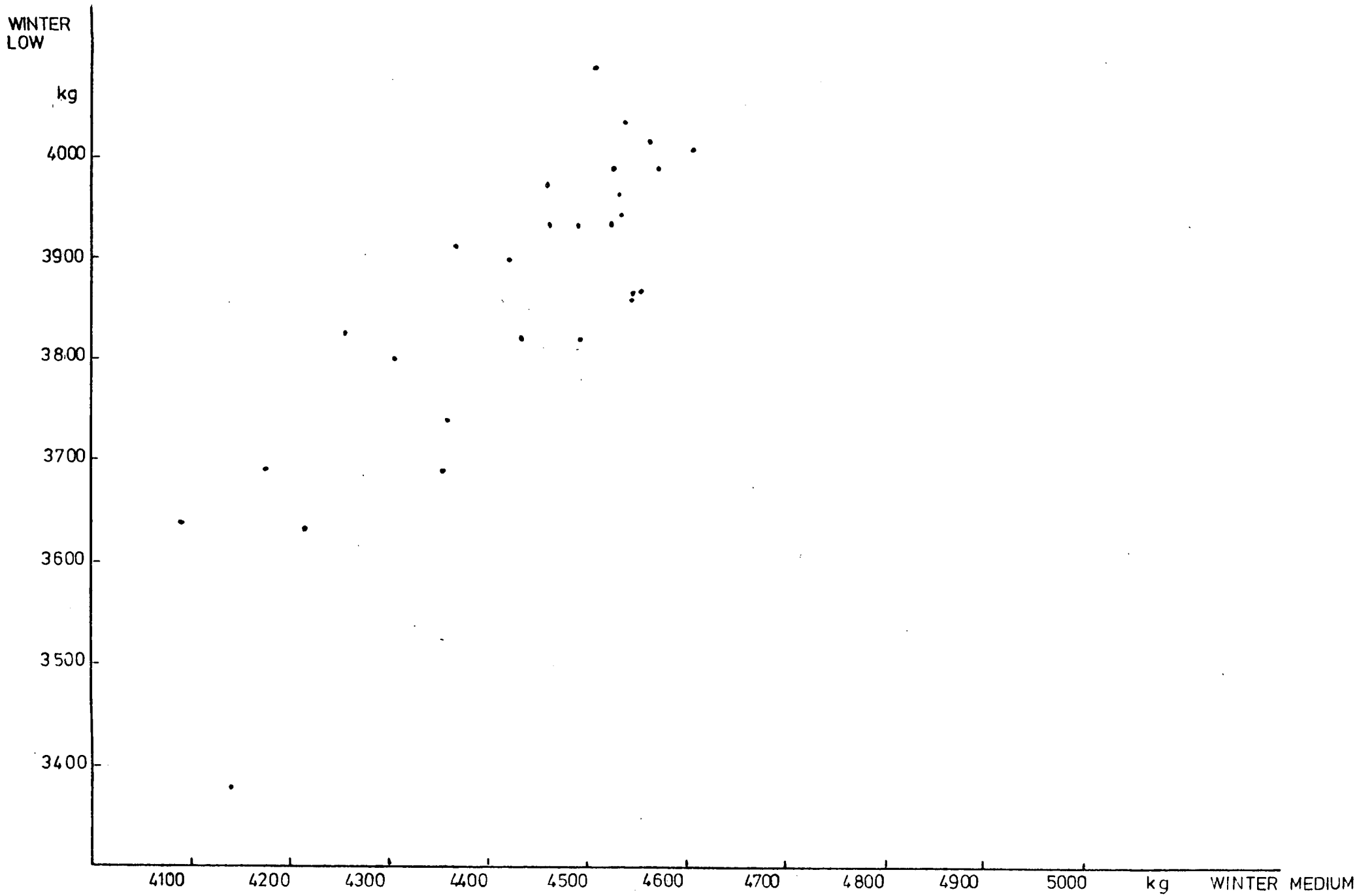


Fig. 2.6 A Scatter Diagram of Least Squares Means of Sires in the Winter Low and Winter Medium Herds (Milk Yield).

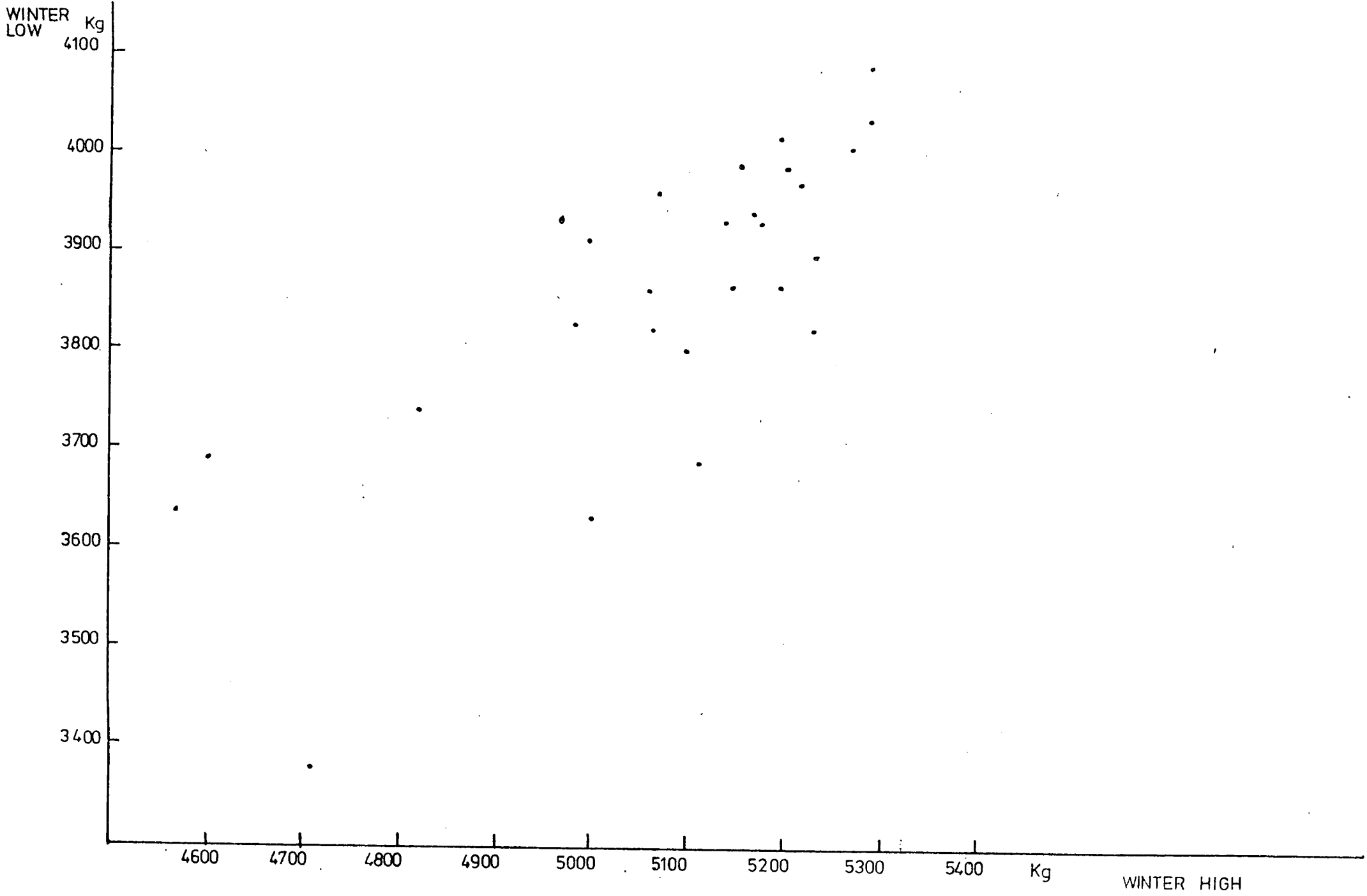


Fig. 2.7 A Scatter Diagram of Least Squares Means of Sires in Winter Low and Winter High Herds (Milk Yield).

C. Sire X Size of Herd Interactions

In this analysis herds are divided into three classes; small, medium and large according to the average number of heifers recorded over four years. A version of model (2.1) without the within sire regressions was fitted to each class separately and three independent estimates of each sires' least-squares means were obtained. The results of this analysis are shown in table (2.21). The within sire mean square showed a slight drop from the large to the small herds. This drop in the within sire variance is given added weight by the drop in the between sires mean square. Given that the average number of effective numbers of daughters per sire in the three levels are not widely different, there is likely to be a real drop in the between sire variance in the large herds and a rise in the within sire variance. When herds were classified using other criteria such as the level of herd production it was consistently noted that a rise in both the within and between sire variance was evident, the latter increasing relatively more. The herd size classification is the only instance where the within and between sire variances have changed in different directions. Although it is difficult to draw conclusions about heritabilities from this set of data, there is a likelihood that the heritability of milk yield is lower in large herds. Fat yield and protein yield exhibit a similar trend in variance over levels while fat per cent and protein yield showed similar but smaller changes in variance.

Table (2.21) shows the analysis of variance carried out on the estimates of sires least-squares means using model (2.10). It is clear that for all traits the sire X level interaction mean

TABLE (2.21)

Analysis Of Variance Table: Between and Within Sires Mean Squares Obtained By Fitting Model (2.1) To Separate Herd Size Classes.

	Small				Medium				Large			
	Error		Sires		Error		Sires		Error		Sires	
	D.F	M.S	M.S	F	D.F	M.S	M.S.	F.	D.F	M.S	M.S.	F
Milk Yield	8566	338298	2937221	7.6	6860	350469	2619949	6.9	7740	379859	2468234	5.6
Fat Per cent	8566	0.0763	1.3256	16.0	6860	0.0765	1.0746	12.9	7740	0.0774	0.7580	9.5
Fat Yield	8566	462.52	3625.11	7.0	6860	460.29	3813.18	7.6	7740	528.14	3438.63	5.7
Protein Per cent	5210	0.0249	0.4697	17.9	4436	0.0263	0.4047	15.1	5195	0.0248	0.3252	11.1
Protein Yield	5210	296.00	1514.31	4.7	4436	303.49	1631.36	5.1	5195	334.72	1314.07	3.3

TABLE (2.22)

Product Moment Correlations Between
Sire Least-Squares Means In
Different Herd Size Classes

Trait		Herd Class	
		Medium	Large
Milk yield	Small	0.8950	0.8582
	Medium		0.7790
Fat per cent	Small	0.9374	0.9132
	Medium		0.9304
Fat yield	Small	0.8927	0.8969
	Medium		0.8194
Protein per cent	Small	0.9313	0.9526
	Medium		0.9307
Protein yield	Small	0.8402	0.8090
	Medium		0.7209

square is not significant. This is emphasised by the product moment correlations shown in table (2.22). Correlations for yield traits are generally high, the lowest correlations being between the medium and small classes of herds. The correlations for percentage traits are all above 0.90. It is likely that for all traits sire X size of herd interaction is close to zero, despite the fact that there are changes in variance over levels.

TABLE (2.23)

Analysis Of Variance Table: Mean-Squares obtained by Using Least-Squares Means Estimates As Observations and Fitting Model (2.10)

		Milk Yield		Fat %		Fat Yield		Protein %		Protein Yield	
	D.F	M.S	F	M.S	F	M.S	F	M.S	F	M.S	F
Reduction due to the mean	1	23276572	739643	20.1900	288428	23266.76	543616	0.0064	18.29	13615.52	3115.68
Size class	2	192333	61.12	0.0002	0.29	297.42	69.49	0.0013	3.71	297.82	68.15
Sires	26	56931	18.09	0.0168	24.00	79.84	18.65	0.0127	36.29	41.05	9.39
Size class X sires	52	3352	1.07	0.0005	0.71	3.85	0.90	0.0003	0.86	3.53	0.81
Remainder	-	3147	-	0.0007	-	4.28	-	0.00035	-	4.37	-

Interactions Involving Age Effects

1. Sire X Age of Calving Interactions

In the analysis of variance resulting from model (2.1) on the original two sets of data which is shown in tables (2.8 - 2.10), linear and quadratic regression on age at calving within sires are fitted. In table (2.8) the linear regressions within sires for both milk yield and fat per cent are significant at the 5% level, while the quadratic regressions are not. The actual estimates of linear and quadratic regressions for each sire are shown in table (2.24). The linear term varies from 14.5 kg per month of age for sire 3 to 81.7 kg per month of age for sire 27. However sire 3 has only 53 daughters calving at intermediate ages. The majority of its daughters (668) calve after 33 months of age. Most of the sires' linear regressions estimates lie between 40 and 50 kg per month of age although clearly there are some sires which fall well short of this. Table (2.25) shows the results of applying model (2.1) after the data are transformed to logarithms. All estimates in this table are given as multiples of 10^{+6}. The estimates vary between 1460×10^{-6} and 7695×10^{-6}. However the latter estimate is obtained from sire 27 which is the sire with the lowest number of daughters in the set. The quadratic regressions on both the linear and the multiplicative scale are insignificant in the sense that they are not different from the single pooled estimate of regression fitted across all levels.

2. Age X Level of Herd Production Interaction

Pooled and within level regressions on age at calving were fitted using model (2.2). The results of the analysis of variance

TABLE (2.24)

Sires' Regression on Age at Calving Esquations for Milk Yield (additive Model)

Sire	Linear				Quadratic			
	Constant Estimate	S.E.	Least-Squares mean	S.E.	Constant Estimate	S.E.	Least-Squares mean	S.E.
Pooled	44.02	2.10	44.02	2.10	-1.17	0.39	-1.17	0.39
1	5.49	6.58	49.52	6.49	0.32	1.20	-0.85	1.18
2	-10.34	6.36	33.68	6.24	-0.25	1.18	-1.43	1.16
3	-29.51	10.22	14.51*	10.39	0.59	2.04	-0.58	2.09
4	2.84	11.69	46.86	11.96	-0.72	2.24	-1.90	2.29
5	6.41	11.33	50.43	11.58	4.28	2.08	3.10	2.13
6	17.94	8.90	61.97	8.99	2.81	1.56	1.64	1.57
7	-18.97	7.15	25.05	7.11	-1.25	1.32	-2.43	1.31
8	-2.28	8.07	41.75	8.11	-0.39	1.45	-1.56	1.46
9	-18.37	14.71	25.65	15.14	-2.23	2.66	-3.40	2.73
10	7.91	10.72	51.93	10.92	2.68	1.98	1.51	2.01
11	-12.77	10.13	31.26	10.31	-1.79	1.64	-2.96	1.66
12	-3.79	12.95	40.23	13.28	-0.83	2.38	-2.01	2.44
13	-13.43	8.48	30.59	8.53	2.64	1.65	1.47	1.66
14	-9.79	7.37	34.23	7.35	-0.45	1.42	-1.63	1.42
15	18.40	14.37	62.42	14.78	-3.97	2.43	-5.14	2.49
16	2.64	11.10	46.67	11.33	-2.22	2.00	-3.40	2.04
17	5.18	9.05	49.20	9.15	0.34	1.61	-0.84	1.63
18	-4.54	9.14	39.48	9.24	0.37	1.96	-0.80	2.00
19	-2.22	9.28	41.80	9.40	-0.52	1.86	-1.69	1.89
20	1.74	13.04	45.77	13.39	-0.91	2.41	-2.09	2.47
21	3.66	8.52	47.69	8.59	1.27	1.57	0.09	1.59
22	8.30	7.16	52.32	7.12	1.51	1.32	0.34	1.32
23	-15.17	12.56	28.86	12.86	-1.49	2.55	-2.67	2.62
24	4.20	12.82	48.23	13.15	0.31	2.26	-0.86	2.31
25	7.05	8.58	51.07	8.65	-0.20	1.71	-1.37	1.74
26	11.72	12.71	55.74	13.03	-0.79	2.11	-1.96	2.16
27	37.70	16.08	81.72*	16.56	0.89	3.00	-0.28	3.09

TABLE (2.25)

Sires' Regression on Age at Calving Least-Squares means ($x10^{-6}$) for milk yield
(Multiplicative Model)

Source	Linear		Quadratic	
	Least-Squares mean	Standard Error	Least Squares mean	Standard Error
μ	3637799	1307		
Pooled	4413	208	-146	38
Sire 1	485	643	-134	117
2	3339	619	-129	115
3	1460	1030	-73	207
4	4656	1185	-308	227
5	4718	1148	280	211
6	6173	891	114	156
7	2206	705	-238	130
8	4308	803	-221	144
9	2936	1500	-373	271
10	4975	1083	122	200
11	3412	1021	-244	164
12	3978	1316	-258	241
13	2726	846	56	165
14	3460	728	-147	141
15	6472	1465	-580	247
16	4480	1123	-345	202
17	4934	907	-102	161
18	3701	916	-125	198
19	4340	932	-197	188
20	5143	1327	-347	245
21	4717	851	2	157
22	4908	706	-41	130
23	3038	1275	-391	259
24	5286	1304	-36	229
25	4910	858	-133	172
26	6330	1292	-220	214
27	7695	1642	124	306

* All values are given as multiples of 10^{+6}.

on both the linear and multiplicative scale are shown in tables (2.11 - 2.13). On the additive scale neither the linear nor the quadratic within level regressions are signficant. On the multiplicative scale, however, the linear within sire regression is significant at 5%. The F-ratio for the regression showed a rise from 0.45 to 3.29 while the ratios for other interaction terms in the model remained the same. It is difficult to interpret this result since the change of scale could have resulted in these interactions while, on the other hand, if the effects of age truly act multiplicatively with the rise in the herd level of production, the level X age interactions would then be genuine. Tables (2.26) and (2.27) show the actual estimates of pooled and within level regressions obtained on both scales. On the additive scales the linear regressions range between 44 and 40 with the highest effect of age being in the medium level. On the multiplicative scale they range between 4613×10^{-6} and 3539×10^{-6} with the highest effect of age in the low level of herd production. Again no useful interpretation can be made of the fact that the ranking of levels changed on the two scales but it is useful to note that on both scales the differences in slope, although statistically significant on the multiplicative scale, are not large.

3. Age X Season and other Interactions

The analysis of variance resulting from combining models is shown on tables (2.28) and (2.29). Age X season interaction estimates have proved to be insignificant for all five traits. The age X season subclasses are large, both age and seasons being

TABLE (2.26)

Levels Regression on Age Equations for Milk Yield (Additive Model)

Source	Constant estimate	Standard error	Least-Squares mean	Standard error
μ	4435.15	8.80	4435.15	8.80
Linear	42.24	1.83	42.24	1.83
Linear within level 1	-0.55	2.66	41.70	3.36
Linear within level 2	2.24	2.49	44.49	2.96
Linear within level 3	-1.70	2.57	40.55	3.15
Quadratic	-0.91	0.34	-0.91	0.34
Quad. within level 1	-0.01	0.49	-0.92	0.61
Quad. within level 2	-0.06	0.46	-0.96	0.54
Quad. within level 3	0.06	0.49	-0.85	0.61

TABLE (2.27)

Pooled and Within Sire least square means ($x10^{-6}$)
on Age at Calving for Milk Yield (multiplication model)

Source	Linear		Quadratic	
	Least-squares mean	Standard error	Least squares mean	Standard error
μ	3637740	886		
Pooled	4190	182	-115	34
Within level 1	4613	333	-126	60
Within level 2	4417	293	-112	54
Within level 3	3539	312	-107	60

* All values are given as multiples of 10^{+6}

TABLE (2.28)

Analysis of variance Table: Pooled Estimates of Mean Squares obtained by fitting Models (28)

	Milk Production			Fat Per Cent			Fat Yield		
Source	D.F	M.S	F.	D.F	M.S	F.	D.F	M.S	F
Herd-year-sire	14567	1358569	3.86**	14567	0.2080	2.75**	9523	2177.09	4.45**
Season	2	25353838	71.97**	2	0.5216	6.89**	2	29602.29	60.57**
Age	2	64389097	182.78**	2	0.2358	3.11**	2	66647.67	136.38**
Herd-year-sire X season	4929	421416	1.20**	4929	0.0920	1.22**	3172	604.37	1.24**
Herd-year-sire X age	5599	398500	1.13**	5599	0.0830	1.10**	3651	560.55	1.15**
Age X Season	4	292577	0.83	4	0.1091	1.44	4	234.02	0.47
Remainder	18727	352268	-	18727	0.0757	-	11943	488.70	-

TABLE (2.29)

Analysis of Variance Table: Pooled Estimates of Mean Squares obtained by fitting Models (2.2)

Source	Protein %			Protein Yield		
	D.F	M.S.	F.	D.F	M.S	F.
Herd-year-sire	9523	0.0645	2.59**	9523	1461.51	4.73**
Season	2	0.4050	16.27**	2	11046.85	35.76**
Age	2	0.0527	2.12	2	46236.38	149.66**
Herd-year-sire X season	3172	0.0308	1.24**	3172	400.40	1.30**
Herd-year-sire X age	3651	0.0287	1.15**	3651	364.96	1.18**
Age X Season	4	0.048	1.94	4	187.33	0.61
Remainder	11943	0.0249	-	11943	308.94	-

divided into three classes giving a total of 9 subclasses for the interaction with an average of 4745 heifers each. With such large numbers of heifers involved it appears that age X season interaction is either non existent, or is close to zero.

Combining models (2.3 - 2.7) also resulted in estimates of a herd-year-sire X season interaction. Both of these proved highly significant and were estimated with large numbers of degrees o freedom. However, the F-ratios for both estimates compared with other terms in the model are very close to one and the numbers of heifers within each interaction subclass is bound to be very small. The total number of degrees of freedom in this analysis is 43830 while it should be 42700. This is because the two estimates of interaction are confounded with each other.

Section 5 DISCUSSION

Introduction

Genotype X environment interactions among wide spread use sires could arise from a variety of sources. Apart from the possibility of a true change of ranking of genotypes over environments such interactions may result from non-random or compensatory mating of sires with certain qualities to cows lacking them. Since this set of sires is a proven one a farmer would know in advance the strengths of each sire and the weaknesses among his cows and mate them accordingly. Another likely reason for possible interactions among wide-spread use sires is the preferential treatment of daughters of particular sires. Besides the possibility of true genotype X environment interactions, both non-random mating and preferential treatment of sire groups seem to apply particularly to wide-spread use sires and since these contribute a large proportion of the dairy population it is important to estimate the magnitude of genotype X environment interactions among them.

Sire X Herd-year-season Interactions

Tables (2.8 - 2.11) show that the sire X Herd-year-season interaction mean square is significant at the 5% level in the case of fat percent, protein percent and protein yield. In fact for protein yield the importance of the interaction mean square is increased significantly after transformation to a logarithmic scale. However, with the exception of protein yield it appears that the relative magnitude of the interaction component is small. This is similar to the conclusions of Bereskin (1963),

Hickman and Henderson (1955) and Legates et al (1956). However, Specht and McGilliard (1960) reported that the sire X Herd interaction component amounted to 7% of the total variance. Probably the most comprehensive study on the subject was made by Kelleher et al (1967). They analysed 37701 Holstein-Friesian records obtained from a wide geographic area in the U.S.A. and spread over 6 herd-year-seasons. To avoid the problem of non-orthogonality in such data, subsets of data were found where orthogonality could be achieved and sums of squares were pooled over all subsets. The interaction component was found to amount to 2% of the total variance and the 95% confidence limit included zero. The authors concluded that there is no reason to believe that the interaction is real and that if it existed it would not bias progeny tests in an appreciable manner.

With regard to this study it is difficult to ascertain whether the sire X herd-year-season interaction really exists since a large percentage of the sire X herd-year-season subclasses are empty. The fact that the size of each of the 'environments' is small and that there is large number of them allows small variations between herds in ranking and variance to appear of significant importance. However, since the relative magnitude of the interaction mean square is small except in the case of protein yield, it is comforting to be able to ignore it since, if it really existed in any significant amount, it is difficult to foresee what could have been done about it in terms of a practical breeding policy because of the very small nature of the environments under study.

Significance Tests and Estimates of Sires' Least Squares Means in Different Levels of Production

The cross-classification analysis of sires X levels of production using model (2.2) given in tables (2.11 - 2.13) shows a significant (5%) sire X level interaction for milk yield and fat yield. In both cases the interaction mean square is slightly higher than the magnitude of the error variation. This situation changes slightly with the change to the logarithmic scale. The product moment correlations shown in table (2.14) indicate that the main source of the interaction is the high X low level classification where the correlation is the lowest (0.77). This relationship is further illustrated in figure (2.1) where the estimates of the least squares means in the low level are plotted against those in the high. These estimates are obtained from an analysis using model 1 on each level of production separately. Most sires are fairly well packed around the regression line apart from seven sires that appear to be exhibiting rather large changes in rank from one level to the other. The numbers of effective daughters and the constant estimates and their standard errors for those seven sires are shown in table (2.15). None of them seems to have particularly low numbers of daughters except sire 15 which had only 13 effective daughters in the high level. The only possible conclusion one can draw from this is that with respect to the high-low relationship some sires such as sire 10 and 9 are for one reason or another genuinely exhibiting genotype X environment interactions of considerable magnitude, although averaged over all sires, this interaction seems to be small. The analysis of variance carried out within each level

separately (Table 2.16)) shows an increase in the within sire variance. If a function of the effective numbers of daughters for all sires is used to approximate the coefficient of variance component in between sires, an increase in the between sire variance component relative to the total variance can be demonstrated.

McDaniel and Corley (1967) carried out a similar investigation on 40 bulls with a minimum of 1000 daughters each and divided into four herd-mate levels of production. Correlations between sire progeny averages in different levels were found to be high ranging from 0.88 to 0.96 with the lowest correlations being between the two extreme classes. As in this study a few progeny groups were found to deviate substantially from the overall pattern but the conclusion was that regardless of the herd-mate production level the bulls will rank the same. There is some disagreement in the literature about whether there is a trend for higher heritability at higher levels of production. However, the weight of the evidence seems to favour the existence of such a trend (e.g. Mason and Robertson 1956, Van Vleck, 1963). It is not possible to estimate exactly the changes in variance over levels in this study since the set of sires under study cannot be assumed to be random. The estimates of product moment correlations and significance tests obtained seem to agree closely with the findings of the authors sited above and we can safely assume that apart from a few exceptions, the majority of bulls will rank similarly in different levels of production.

Another method by which sires X herd production level was studied was by dividing herds into two groups of winter and spring

calving herds. The winter calving herds were further divided into three levels of production; low, medium and high. The differences in the mode of calving between herds reflect regional and managerial differences. Spring calving herds tend to be dominant in the south where the milder weather allows the herd managers to keep cattle more often out-doors and to make the best use of available pastures the heifers are usually calved in spring. The farms tend to be of larger size than the average for the United Kingdom and generally the management practices tend to reflect the characters of an extended system of production.

Tables (2.18 and 2.19) show the results of the analysis carried out using sires' least square means estimates as observations according to model (2.9). In the case of milk yield the season X sires interaction mean square is significant at 10% while the levels within season X sires interaction was significant at 5%. In order to judge the importance of this interaction the estimes of least squares means obtained in each level are plotted against each other in figures (2.4 - 2.7). It is apparent that, as was found previously, the majority of sires exhibit no substantial genotype X environment interactions. However the same group of bulls as before, deviate substantially from the regression line. As expected the lowest correlation in milk yield was between the high winter and spring (0.70). In the case of fat %, fat yield and protein % the correlations are all reasonably high, the lowest correlation between any two classes being 0.72. In the previous analysis where herds were classified into low, medium and high levels of production the correlations of protein yield over different levels were lower

than those for other traits with a minimum of 0.64 and a maximum of 0.80. This state of affairs holds well when winter herds are divided into low, medium and high. The minimum correlation between these classes is 0.61 and the maximum is 0.74. However, the correlations of all these classes with the spring herds are very low indeed. They range between 0.28 and 0.52. Clearly on the basis of an average classification of herds into three levels the correlations for protein yield are on the low side, but nevertheless, are within the range of chance deviations from previous estimates. This correspondence disappears when the spring class of herds is introduced into the picture. This is surprising since, given the reported high correlation between protein yield and milk yield, one would expect that the estimates for the two would be close. Because of the nature of this result the data were screened for outlying values and the method of computing protein yield from protein % and milk yield was revised to eliminate rounding errors. The result held well and, as can be seen from figure (2.4), it is clear that the interactions are not a result of a small group of sires with large deviations, but rather an extensive change of ranking between winter and spring herds. For all yield traits the correlations involving the spring classification of herds are the lowest and this probably reflects the fact that the winter and spring classification reflects to a greater degree than other classifications, managerial and regional differences between herds. The estimates of correlation for milk yield were also estimated from the protein subset and they proved to be generally lower than those obtained from the milk and fat subset which indicates a set effect that may partially account for the drop in the estimates of correlation for protein yield.

This problem will be taken up again in later chapters when estimates of the heritability of protein yield in the two levels are obtained.

The herd level of production was also defined in terms of herd size classes. The change in herd size is probably the most radical change in the British dairy industry in recent years. Traditionally the Southern, Midwestern and South Eastern regions are large herd areas where over a third of the herds were of 50 cows or more and 10% are herds of at least 80 cows in 1970 (MMB Dairy herd census, 1970 and the report of the Breeding and Production Organisation 1973/74). The Eastern and Midlands regions are also dominated by large herds and small herds are more dominant in the Welsh and far Western regions. This indicates that besides reflecting differences in management and the degree of personal attention that a herd owner can afford, herd size classes may also reflect regional differences and thus important environmental differences may be expected to exist between different herd size classes.

Table (2.22) shows the product moment correlations between sires' least squares means in the three herd size classes. For most traits the lowest correlations are those between the medium and large classes of herds and protein yield exhibited the lowest correlation of 0.72 between large and medium classes. However, as can be seen from table (2.17), genotype X environment interactions for all traits are not significant and there is a high correspondence between estimates of sires over the three size classes.

Interactions Involving Age Effects

Adjustment of lactation milk yield records for the age of the cow at calving has the major purpose of facilitating comparisons between daughter groups with different age structures. One source of variation that could affect this adjustment is that the variation between sires in the degree of maturity or development of their daughters at a certain age. If sires do vary in the maturity of their daughters, then adjustment by using a common regression or mature equivalent factors could favour the progeny of the sire whose daughters are inherently early maturing.

In this study the sire effect is significant on both the additive and multiplicative model at a probability of 0.05 for both milk yield and fat percent (Tables (2.8, 2.9). This indicates the presence of differences between widely used sires in the rate of maturity of their daughters and that common age correction factors will not fit accurately the progeny of each sire. Table (2.24) shows the regression equations for individual sires using the additive model. Clearly there is a large variation in the regression factors between sires. The linear regression coefficient ranges between 15 $\pm$ 10 and 82 $\pm$ 17. Hillers and Freeman (1965) found that there was a large variation between sires in the rate of maturity to the extent that some sires exhibited negative regressions.

Hargrove (1973) analysed 18869 first lactation records with sires having at least 50 daughters distributed in at least 10 herds in one year. He found a highly significant sire effect ($P<0.01$) and heritabilities of the rate of maturity ranged from 0.085 to 0.108 for fat and from 0.099 to 0.115. The practical implications of these results is that since the effect of the

inequality in the rate of maturity between sires is not removed by the type of correction factors used at present, and as the indications are that the rate of maturity is heritable (Hickman & Henderson (1955)), selection for first lactation production must exert some selection pressure on the rate of maturity. If the rate of maturity and first lactation production are positively correlated then present selection practices favour rapid maturing progeny groups.

Tables (2.11 - 2.13) show the results of fitting linear and quadratic regressions within different levels of production using both an additive and a multiplicative model. Table (2.11) shows that if we assume that age effects affect heifer milk yield additively then that variation is not related to the herd production level. This is also true of all other traits studied. However if we assume that age effects act on milk yield in a multiplicative manner (Table (2.12)) then the amount of age adjustment necessary varies with the herd production level and the ranking of different levels varies from one age group to the other. This is true only of milk yield, other traits being unaffected by the change to the multiplicative model. Both the average and the within level regressions account for relatively more of the variation in milk yield under the multiplicative model.

Searle (1962) suggested that multiplicative factors introduced or increased age X herd interactions. It is certainly the case in this study that age X herd level of production became significant after the change to the multiplicative scale but it is difficult to ascertain if this is a true genotype X environment interaction or a mere artefact of the change in scale.

Figure (2.8) shows a graphic illustration of regression equations obtained by fitting model (2.1) on separate herd classes. The quadratic part of the equation was not significant except in the case of the spring calving class of herds. However the equations used to plot the graphs contain both the linear and quadratic terms to approximate the practice of the MMB. Factors used by the MMB are also shown in the figure. The MMB Factors and the within herd-class equations are plotted as deviations from the population mean. The MMB correction factors exhibit a curvilinearity which is not evident from the results of this study except in the case of the spring calving class of herds, which showed a significant quadratic term. The MMB factors are gross factors that are not adjusted for the level of herd production (Hickman, 1973). For heifers up to 28 months of age the correction factors seem to underestimate the necessary compensation for them, the most affected being young heifers calving in the spring calving class of herds. At twenty two months of age the difference between the MMB correction factors and the spring calving herds deviation estimate is about 125 kg. However, corrections for all class of herds are under-estimated at that age by at least 50 kg. Corrections for heifers between the ages of 28 and 36 appear to be over estimated with those heifers calving in high production winter herds the most disadvantaged. The differences between the three herd classes in the amount of correction needed for age seem to be in the region of 50 - 100 kg at the younger and older ages while the difference is reduced to a minimum over intermediate ages.

The analyses on separate classes of herd using model (2.1) and the cross classification analysis on the multiplicative scale using model (2.2) seem to indicate that the linear within level

2.8 Quadratic regression Curves for first lactation Milk Yield in Separate Herd Classes.

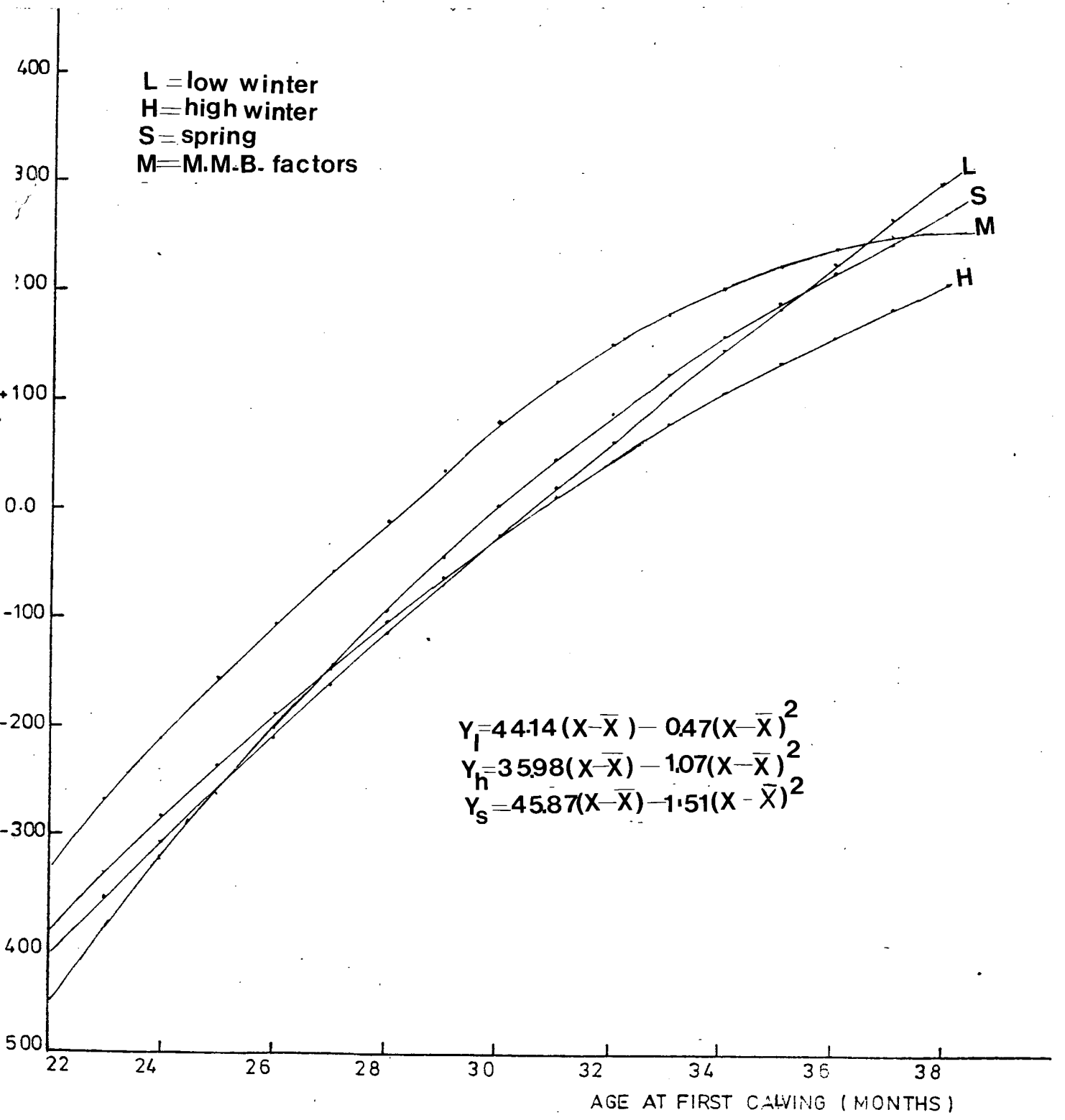


Fig. 2.8

regression on age at calving is significant. The significance of the difference between herd classes seems to be enhanced by the introduction of the winter-spring classification of herds. Hickman and Gravir (1968) found that age correction estimated as the difference between first and later lactation yield is linearly related to herd level of production. Miller (1973) analysed a total of more than three million first and later lactation records made in more than forty thousand herds across the United States. He found that the additive increase in milk yield from 24 months to maturity was substantially larger in high producing herds than in low producing herds in all seasons. He also found that if multiplicative factors are used, adjusting for herd level of production could be ignored due to the uniformity of multiplicative factors over production levels.

In this study age X season interactions were found to be insignificant for all five traits. However, there is a great deal of evidence in the literature pointing to the importance of the season of calving as a source of variation in calving age correction factors (e.g. Lee and Hickman, 1973, McDaniel and Corley, 1966, Gravir and Hickman, 1966). Both classifications of seasons and age in the analysis were fairly large and contained considerable numbers of observations. However, the non-orthogonality of the data was highlighted by the introduction of the herd-year-sire X age and the herd-year-sire X season interactions, both of which proved insignificant. The discrepancy in the total number of degrees of freedom shows that these interactions could not have been computed simultaneously. Both interactions are very small and it is unlikely that sire means in different herd-years exhibit interactions with seasons.

As stated previously the types of environment to which dairy cattle are exposed are fairly uniform. It appears that for all classes of interaction studied, with the possible exception of age X level of herd production interaction, widespread use dairy sires seem to rank similarly over the different classifications of environment studied and that interactions are either small and caused by a few identifiable sires or are close to zero.

Conclusions:

The sire X level of herd production interactions although significant on the multiplicative scale, are probably small. A few sires do exhibit important changes in rank and these same sires have been responsible for interactions detected in other classifications of environment. It is also clear that large changes in the between sires mean squares over different environments occur but, as this set of sires is a highly selected one no firm conclusions can be made about the changes in variance and the problem will be taken up later with a set of random test sires. In the sire X calving season of herd interaction protein yield exhibited very large genotype X environment interactions. The product moment correlations most affected were those involving the spring class of herds which in the protein set had very small numbers of effective daughters for some sires. The large shifts in rank exhibited by those sires with small numbers of daughters may have been responsible for the low estimates of correlation obtained from the protein set. Of the three types of herd classifications used herd size was probably the least important in terms of differences in yield and also showed the least sire X herd level interactions.

With regard to old sires exhibiting genotype X environment interaction it may be useful to establish a simple screening method which is designed to detect such sires and establish a strategy to make the best use of them. Age effects interactions with levels and sires were found to be statistically significant in this analysis, but, they were small compared to other effects. The results of this analysis indicate that no particular action is necessary with regard to age interactions with other environmental or genetic effects.

Chapter 3

Analysis of Records of Daughters of Young Test Bulls

Section 1: Introduction

This analysis was initiated in order to obtain estimates of heritabilities and genetic correlations over low and high levels of herd production. For this purpose the analysis was conducted on a group of young test bulls with much smaller numbers of daughters than those included in the analysis reported in chapter 2. For this reason the set of data used is not particularly useful for the estimation of the genetic correlation but estimates of heritability with reasonably low standard errors can be obtained. Evidence from the previous analysis on widely used bulls indicates that the correlations between the least-squares means of bulls in different levels of production are likely to be high and it remains to determine the changes in the relative magnitude of the genetic variance. This will make it possible to judge whether similar accuracy of progeny testing can be achieved in the low and high level of production herds. The main aim of this analysis will be to obtain estimates of heritability in two levels of production for five dairy traits. The methods of preparation of data for analysis and the method of classification of herds is similar to that explained in chapter 2.

Section 2: Materials

A total of 9795 first lactation records made by daughters of 225 bulls with an average of 43 daughters per sire were analysed. The sires were tested over the three years 1970-72 in four separate regions. Two sets of data were created, the first

(data set 1) comprises records on milk yield, fat per cent and fat yield. The second (data set 2) is made up of all the records in the first set after discarding those with no information on protein per cent or protein yield. Herds were then classified into low and high producing herds depending on the average milk yield of all heifers recorded between 1970 and 1974. In the first data set the low level of production comprised 3078 herd-year-seasons with 5053 daughters of young bulls while the high level included 2862 herd-year-seasons with 4712 daughters of young bulls. The low and the high levels averaged 3800 and 4650 kilogrammes of milk respectively, the dividing point between the two levels being 4150 kilogrammes. Table (3.1) shows the sires and the total number of heifers analysed in each region in the first set of data. Table (3.2) shows the corresponding figures for the second set of data.

A total of 12610 records made by daughters of 84 contemporary old selected sires were used to create connections between different held-year-season classifications. The only criterion for selection of these old sires was the number of contemporary herd-year-seasons in which they were represented. Each region was analysed separately with the South West and South Wales region being analysed in two runs, a run on each year of testing, because of the large numbers of sires involved.

Regions, Numbers of Sires, and Numbers of Records

in the First Set of Data (milk, fat)

TABLE (3.1)

Analysis Number	Region	Test years of sires	No. of Sires	Low Level		High Level	
				No. of heifers	No. of herd year-season	No. of heifers	No. of herd year-seasons
1	North Wales and Midlands	1970/71	42	1030	716	795	610
		1971/72	18	296		313	
		Old	16	777		884	
2	North	1970/71	28	781	547	695	489
		1971/72	17	315		273	
		Old	17	663		902	
3	South West and South Wales	1971/72	32	605	471	642	451
		Old	17	1220		1416	
4	South West and South Wales	1970/71	43	1183	834	887	689
		Old	17	1364		1830	
5	South East	1970/71	32	628	510	855	623
		1971/72	13	215		252	
		Old	20	1339		2215	
	Total		312	10416	3078	11401	2862

Regions, Numbers of Sires, and Numbers of Records

in the Second Set of Data (Protein)

TABLE (3.2)

Analysis Number	Region	Test years	No. of sires	Low Level		High Level	
				No. of heifers	No. of herd year-seasons	No. of heifers	No. of herd year-seasons
1	North Wales and Midlands	1970/71	41	911	619	620	493
		1971/72	17	261		260	
			16	593		710	
2	North	1970/71	28	704	482	602	433
		1971/72	16	280		229	
			17	552		720	
3	South West and South Wales	1971/72	32	571	438	594	412
			17	1135		1294	
4	South West and South Wales	1970/71	43	1120	779	791	634
			17	1292		1737	
5	South East	1970/71	32	585		704	
		1971/72	13	200		215	
			20	1086		1845	

Section 3: Statistical Methods

Statistical Model: The main objective of this analysis is to study the problem of genotype X environment interaction as determined by differences in heritability between different environments and by the genetic correlation coefficient over these environments. As the statistical significance test was considered to be less informative than the above mentioned method it was decided to carry out the analysis within environments rather than over environments. Such design should cater for the possibility of differences in variance within the environments. These differences in variance can often be removed by transforming the data to a different scale (e.g. by using a multiplicative model) but it is not clear whether such a transformation may not create interactions in the usual sense(changes in ranking). A multiplicative model was not used in this study for this reason and also because whatever the result of such a transformation may be the results cannot be interpreted with any certainty. Moreover it is important to investigate the operation of the usual additive genetic and environmental effects in relation to genotype X environment interactions.

A least-squares analysis for data with unequal subclass numbers was made within environments following the method of Harvey (1960). To complete the analysis the following model was used to describe each variable

$$Y_{ijklm} = u + H_i + B_j + S_{jk} + T_l + b_1(X_1 - \bar{X}_1) + b_2 (X_2 - \bar{X}_2) + b_3 (X_2 - \bar{X}_2)^2 + e_{ijklm} \qquad (3.1)$$

where

Y_{ijklm} = the observation made on the mth heifer born in the ith month sired by the Kth sire in the jth batch and in the ith herd-year-season.

μ = the fixed overall mean

H_i = the effect of the ith herd-year-season

B_j = the effect of the jth batch, J = 1, a the ath batch being a set of old sires

S_{jk} = the Kth sire within the jth batch, K = 1, b

T_l = the effect of the lth month nested with seasons l = 1, 4

b_1 = the linear regression coefficient on length of lactation (X_1)

b_2 = the linear regression coefficient on age at calving (X_2)

b_3 = the quadratic regression coefficient on age at calving (X_2).

e_{ijklm} = the random error term associated with the mth individual.

Assumptions about the distribution of effects

Herd-year-seasons effects are deemed to be fixed. Because of the large number of herd-year-season equations, they were absorbed and the analysis completed within herd-year-seasons. The effect of batches is also taken to be fixed. The effect of young sires nested within batches is assumed to be random and normally distributed with variance equal to 6_s^2 while the effect of old sires is taken to be fixed. The error term is assumed to be random and normally distributed with variance $6_e^{2'}$ and all other effects in the model are fixed. The expectations of the

mean squares for the pooled analysis of variance for those parts of the model that are important in the estimation of parameters are shown in table (3.3). The total number of sires is taken to be (a + 8) and from that the number of degrees of freedom lost, because eight separate sets of sires are analysed is subtracted. Ni is the degrees of freedom available for esti-

Table (3.3)

Expectations of the mean squares (E.M.S.)

Source	D.F	E.M.S.
Between sires	a-1	$\sigma_e^2 + K\sigma_s^2$
Within sires	$\sum_{i=1}^{i=8} n_i$	σ_e^2

mating the error variance in the ith analysis, i = 1, 8.

The coefficients of the components of variance for the sets of young sires were computed by a direct method described by Harvey (1960). The programme imposes the restriction that $\sum \hat{S}_{jk} = 0$ and the sums of squares for sires are computed by the matrix multiplication $B^1 Z^{-1} B$, where B^1 is a row vector of a set of constants (for sire effects in this case), Z^{-1} is the inverse of the square symmetrical segment of the variance - covariance inverse corresponding to the set of constants and B is a column vector of the constants. The equation used to compute the K values is as follows:

$$K = \frac{1}{q}\left(\sum_i Z^{ii} - \frac{1}{q-1}\sum_i \sum_j Z^{ij}\right)_{i \neq j} \quad \text{(Harvey 1960)} \qquad (3.2)$$

Where the superscripts on Z specify the particular elements in the matrix inverse corresponding to the square symmetrical segment from the variance covariance matrix, and q is the number

of sires in a group within batches. In simpler terms the first element within the brackets in equation (2) is a summation of diagonal elements of the inverse segment and the second term is the sum of off-diagonal elements divided by the degrees of freedom.

The Estimation of Parameters:

The formula for the estimation of the heritability (h^2) was taken from Falconer (1960):

$$h^2 = \frac{4 \quad 6_s^2}{6_s^2 + 6_w^2} \tag{3.3}$$

The formula for the estimation of the sampling variance of the heritability was also taken from Falconer (1960):

$$6_{h^2}^2 = \frac{32\left[1 + (n-1)t\right]^2 (1-t)^2}{n(n-1)\ (N-1)} \tag{3.4}$$

Where: n = The effective number of daughters per sire

N = The number of sire families

t = The intraclass correlation

In the case of the pooled analysis the term (N-1) was taken to be (N-8) which is the correct number of degrees of freedom and the other terms replaced by their appropriate pooled estimates.

In order to obtain the genetic correlation coefficient ($\hat{r}g$) the covariance of sires least squares means across environments was computed. The expectation of this is:

$$E\ (Cov\ (\bar{X}_1,\ \bar{X}_2)) = Cov_B \tag{3.5}$$

Where $\bar{X}_1$ and $\bar{X}_2$ are the sires' least squares means in the low and high levels respectively, and Cov_B is the genetic covariance across environments. The expectation of this covariance is free from environmental elements since observations were made on

different animals in each environment. This covariance is then divided by the square root of the product of the two between sire variances obtained from the analysis of variance, in order to obtain the genetic correlation ($\hat{r}_g$) i.e.:

$$\hat{r}_g = \frac{\text{Cov}\ (\bar{X}_1,\ \bar{X}_2)}{\sqrt{\sigma^2_{s_1} \times \sigma^2_{s_2}}} \qquad \text{(Robertson, 1959)} \qquad (3.6)$$

The formula for estimating the sampling variance of the genetic correlation was derived by Guiard and Herrendorfer (1977):

$$V(\hat{r}_g) = \frac{r_g^2}{s-1}\left\{\left[\frac{1}{r^2}\left(1 + \frac{1-t_1}{n_1 t_1}\right)\left(1 + \frac{1-t_2}{n_2 n_2}\right) - 1\right](1-r_g^2) + \frac{1}{2}\left(\frac{1-t_1}{n_1 t_1} + \frac{1-t_2}{n_2 t_2}\right)^2\right\} + \frac{r_g^2}{2_s}\left\{\frac{(1-t_1)^2}{n_1^2(n_1-1)t_1^2} + \frac{(1-t_2)^2}{n_2^2\,(n_2-1)t_2^2}\right\} \qquad (3.7)$$

Where s is the number of sire families and the subscripts 1 and 2 refer to estimates in the low and high level respectively. Another estimate of $\hat{r}_g$ was obtained using a weighted covariance analysis. The weights attached to each sire were worked out as follows: Take a sire Si with a least-squares mean $\alpha_{s_{ii}}$ in environment 1 and α_{si2} in environment 2. Its weighting (Ws_i) is:

$$W_{s_i} = \frac{1}{\hat{r}g^2 + \left(1 + \frac{1-t_1}{t_1 n_{s_{i1}}}\right)\left(1 + \frac{1-t_2}{t_2 n_{si2}}\right)} \qquad (3.8)$$

Where n_{si1} and n_{si2} are the numbers of effective daughters in environment 1 and 2 respectively. The covariance within a single set of young sires is then:

$$\text{Cov}\ (\bar{X}_1, \bar{X}_2) = \frac{\Sigma_i \hat{\alpha}_{s_{i1}} \hat{\alpha}_{s_{i2}} W_{si} - \Sigma_i \hat{\alpha}_{si1} W_{si} \Sigma_i \hat{\alpha}_{si2} W_{si}}{\Sigma W_{si} - \bar{W}_h} \qquad (3.9)$$

Where $\bar{W}_h$ is the average weight within the set of young sires. The sum of the numerator and denomenator separately overall sets of young sires provided the weighted estimate of the covariance. This process is repeated by inserting the new value of $(\hat{r}g)$ in equation (6) until a stable value of $(\hat{r}g)$ is obtained.

Methods of Combining analyses on different sets of Young Sires

1. Unweighted Pooling

The different ANOVAS take the following form within each level:

Source	d.f	M.S	E(M.S)
Between sires	b_i	B_i	$6_w^2 + K_i 6_B^2$
Within sires	w_i	W_i	6_w^2

Pooling is done by adding up sums of squares, degrees of freedom and K-values so that the pooled sum of squares between sires (S.S.B.) will be equivalent to $\Sigma b_i B_i$. The expectation for this pooled sum of squares is

$$E(S.S.B.) = \Sigma b_i\ 6_w^2 + \Sigma b_c K_i 6_B^2 \qquad (3.10)$$

The pooled mean squares between sires (M.S.B.) is then calculated as $\frac{\sum b_i B_i}{\sum b_i}$ and the expected value for this pooled mean square is:

$$E(M.S.B.) = \sigma_w^2 + \frac{\sum b_i K_i \sigma_B^2}{\sum b_i} \qquad (3.11)$$

The unweighted pooling of covariances was done in a similar manner so that if $\bar{X}_i^*$ is its least squares mean in the other level then the pooled covariance is equivalent to:

$\frac{\sum b_i Cov(\bar{X}_i, \bar{X}_i^*)}{\sum b_i}$ The expected value for this quantity is the genetic covariance between sires (CovB). Weighted pooling for both mean squares and covariance analysis over the five runs was done in the manner described by Nelder and Wedderburn (1972).

Section 4: Results

The analysis has been carried out on five separate runs following the natural split of the data into regions. In all five runs a set of old sires was included and in three of them two sets of young sires were analysed simultaneously. The decision of whether to analyse different sets of young bulls together was made solely on the basis of the programme capacity. Differences between old and young sets of sires were systematically high and correlations between the least-squares means of old sires in the low and high levels of production ranged between 0.7 and 0.9. The results of the analysis on young sires will be presented in three steps: First the results of the analysis on the first set of data that comprises milk yield, fat percent and fat yield. Then the results on the second data set will be presented and these include mainly the results of protein percent. and protein yield analysis, however, the results on milk yield, fat percent and fat yield are also included for the purpose of set comparisons. Thirdly, the combined analysis on the first and second data sets will be presented.

Table (3.4) shows the effective numbers of daughters, least-squares means, and standard errors for all five traits in the five separate analyses. These means are obtained over both young and old sires and tend to show a consistent difference of about 800 kilogrammes of milk between the low and high levels of production. The number of effective daughters in the low levels is 4635 and in the high level it is 4711.

Table (3.4)

Effective Numbers of Daughters, Least-Squares Means and Standard Errors obtained From the Five Separate Analyses

Trait	Analysis Number	Low			High		
		Effective Number	Least-Squares mean	Standard errors	Effective Number	Least-Squares mean	Standard error
Milk yield	1	987	3793.18	34.98	888	4690.11	45.22
	2	898	3796.39	39.14	733	4708.59	49.97
	3	867	3903.90	45.11	836	4642.45	47.68
	4	1098	3783.70	17.70	1235	4627.13	39.18
	5	785	3826.34	54.26	1019	4610.86	47.71
Fat percent	1	987	3.7609	0.0164	888	3.8422	0.0186
	2	898	3.7555	0.0187	733	3.7960	0.0217
	3	867	3.7584	0.0199	836	3.8210	0.0197
	4	1098	3.7699	0.0148	1235	3.8306	0.0149
	5	785	3.7695	0.0243	1019	3.7307	0.0214
Fat yield	1	987	142.42	1.42	888	179.65	1.89
	2	898	142.28	1.70	733	178.20	2.15
	3	867	146.13	1.87	836	176.76	2.00
	4	1098	140.57	1.65	1235	177.41	2.43
	5	785	144.05	2.26	1019	171.25	1.90
Protein percent	1	857	3.1906	0.0098	751	3.2308	0.0105
	2	745	3.2045	0.0103	614	3.2305	0.0126
	3	811	3.1812	0.0112	773	3.2120	0.0105
	4	1027	3.1739	0.0084	1133	3.2206	0.0086
	5	672	3.2312	0.0134	878	3.2645	0.0111
Protein yield	1	857	121.12	1.30	750	151.20	1.64
	2	745	121.24	1.48	614	150.75	1.70
	3	811	123.66	1.68	773	148.13	1.52
	4	1027	119.56	0.73	1133	148.41	1.44
	5	672	125.48	1.93	878	149.92	1.75

In the second set of data (the protein set) these numbers are reduced by about 500 effective daughters in each level.

Table (3.5) shows the degrees of freedom and K-values for the first set of data. The results of the analysis of variance on the same data set are shown in table (3.6). Estimates of the within sire variance, between sire variance and heritabilities are given. Analysis on protein percent and protein yield was excluded in this data set because many records did not have information on either. The noise caused by the large standard errors in the separate analyses is evident. Estimates of heritability of milk yield in the low level vary between 0.01 and 0.40. In the high level they vary between 0.15 and 0.38. However, in six out of the eight sets of young sires analysed there was a consistent trend for higher heritabilities at the higher level of production in yield traits due to a relatively greater increase in the genetic compared to the within sire variance.

The degrees of freedom for sires and error mean squares and the K-values for the second set of data are shown in table (3.7). It can be seen that there is a drop in the numbers of effective daughters in the order of two effective daughters per sire. However, the set of 16 sires in the first analysis showed a rise in the number of effective daughters per sire probably because of the loss of a number of herd-year-season X sire subclasses with low numbers of records. Generally it is worth noting how many effective daughters are lost in correcting for environmental variation. Originally the number of effective daughters was 40 daughters

Table (3.5)

Degrees of Freedom for Sires and Error Mean Squares and K-Values in the first Data Set

Analysis Number	Low			High		
	Error	Sires	K	Error	Sires	K
1	1300	41	10.88	1295	41	8.45
		17	4.37		17	4.59
2	1139	27	14.76	1308	27	14.06
		16	9.89		16	8.76
3	1294	31	10.64	1547	31	12.49
4	1642	42	13.13	1957	42	11.30
5	1596	31	11.34	2517	31	16.17
		12	7.70		12	10.70

Table (3.6)

Estimates of Heritability, Within Sire and Between Sire Variance for Data Set 1

Analysis Number	Trait	Level 1 (low)			Level 2 (high)		
		$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$
1 (Batch 1)	Milk yield	250841.30	28195.27	0.4042	366937.74	21646.46	0.2228
	Fat %	0.0749	0.0101	0.4753	0.0780	0.0197	0.8166
	Fat yield	364.68	28.58	0.29	541.56	43.91	0.3000
1 (Batch 2)	Milk yield	250841.30	354.07	0.0056	366937.74	34000.52	0.3392
	Fat %	0.0749	-0.0021	-0.1154	0.0780	0.0015	0.0755
	Fat yield	364.68	25.63	0.2627	542.20	48.58	0.3289
2 (Batch 1)	Milk yield	251738.66	14253.42	0.2143	399016.21	41911.48	0.3802
	Fat %	0.0713	0.0117	0.5643	0.0812	0.0065	0.2972
	Fat yield	346.02	18.58	0.2038	568.44	24.58	0.1658
2 (Batch 2)	Milk yield	251738.66	07162.92	0.1107	399016.21	29896.89	0.2788
	Fat %	0.0713	0.0005	0.0266	0.0812	0.0045	0.2112
	Fat yield	346.02	10.24	0.1150	568.44	1.92	0.0134
3	Milk yield	316950.65	4194.17	0.0520	383213.58	31162.50	0.3008
	Fat %	0.0781	0.0097	0.4433	0.0759	0.0051	0.2508

Analysis Number	Trait	Level 1 (low)			Level 2 (high)		
		$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$
3	Fat yield	379.49	-76.01	-1.0018	526.53	47.68	0.3321
4	Milk yield	264164.75	14730.77	0.2113	378802.27	39717.33	0.3796
	Fat %	0.0694	0.0077	0.3996	0.0741	0.0048	0.2448
	Fat yield	342.42	13.99	0.1570	512.59	34.74	0.2539
5 (Batch 1)	Milk yield	286935.58	20273.42	0.2640	399801.17	15733.37	0.1515
	Fat %	0.0775	0.0031	0.1517	0.0883	0.0095	0.3887
	Fat yield	414.48	9.88	0.0931	535.76	14.24	0.1036
5 (Batch 2)	Milk yield	286935.58	26453.05	0.3376	399801.17	15712.68	0.1513
	Fat %	0.0775	0.0008	0.0395	0.0883	0.0043	-0.2059
	Fat yield	414.48	12.59	0.1179	535.76	38.26	0.2666

Table (3.7)

Degrees of Freedom for Sires and Error Mean Squares and K-Values in the Second Data Set

Analysis Number	Low			High		
	Error	Sires	K	Error	Sires	K
5	1061	40	9.96	1012	40	6.36
		16	6.60		16	6.07
	982	27	13.50	1046	27	12.00
		15	9.33		15	7.62
	1208	31	10.01	1416	31	11.66
	1562	42	12.57	1823	42	10.30
	1346	31	10.52	2065	31	13.34
		12	7.40		12	8.17

per sire and that drops to about 11 daughters per sire after correcting for herd-year-seasons effects. Three young sires were also lost in this set because they did not satisfy the condition that each sire should have at least 5 daughters in each level after the records with no information on either protein percent or protein yield were discarded. Table (3.8) shows the results of the analysis of variance on the second set of data. In addition to the three traits whose results were given in table (3.6), this table also shows the results on protein percent and protein yield. Estimates of heritability of the three traits, milk yield, fat percent and fat yield are close to those obtained using the first data set although generally lower.

Table (3.9) shows a summary of the pooled analysis of variance comprising the degrees of freedom and the between and within sire mean squares for the five traits. Table (3.10) shows the estimates of the pooled within sire and between sire variances from the two sets of data. It is evident that, overall, milk yield has shown an increase in the between sire variance relative to the total variance in both sets of data. This relative increase is smaller in the second set of data. A similar trend can be demonstrated in the case of fat yield for which the relative amount of between sire variance has doubled in the high compared to the low level. Percentage traits did not show a significant change in the relative magnitude of the between sire variation. Table (3.11) shows

Table (3.8)

Estimates of Heritability, Within Sire and Between Sire Variance for Data Set 2

Analysis Number	Trait	Level 1 (low)			Level 2 (high)		
		$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$
1 (Batch 1)	Milk yield	254456.57	32190.26	0.4492	356987.97	11558.19	0.1254
	Fat %	0.0757	0.0098	0.4594	0.0797	0.0186	0.7556
	Fat yield	371.26	33.98	0.3354	534.61	33.85	0.2382
	Protein percent	0.0269	0.0051	0.6334	0.0255	0.0058	0.7417
	Protein yield	240.99	28.10	0.4177	301.98	17.61	0.2205
1 * (Batch 2)	Milk yield	254456.57	-7122.47	-0.1150	356987.97	20601.43	0.2182
*	Fat %	0.0757	-0.0009	-0.0492	0.0797	-0.0014	-0.0724
	Fat yield	371.26	5.96	0.0632	534.61	26.32	0.1877
	Protein percent	0.0269	0.0000	0.0045	0.0255	0.0064	0.2009
*	Protein yield	240.99	-10.48	-0.1819	301.98	38.97	0.4572
2 (Batch 1)	Milk yield	262285.54	13315.74	0.1933	401609.01	49271.13	0.4371
	Fat %	0.0717	0.0127	0.6033	0.0806	0.0066	0.3044
	Fat yield	356.88	18.40	0.1961	575.76	21.07	0.1412

Analysis Number	Trait	Level 1 (low)			Level 2 (high)		
		$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$
2 (Batch 1)	Protein percent	0.0268	0.0034	0.4510	0.0256	0.0035	0.4813
	Protein yield	237.58	11.46	0.1840	366.79	37.19	0.3683
2 (Batch 2)	Milk yield	262285.54	-3824,04	-0.0592	401609.01	45672.69	0.4084
	Fat %	0.0717	0.0004	0.0233	0.0806	0.0029	0.1393
	Fat yield	356.88	5.87	0.0647	575.76	4.10	0.0283
	Protein percent	0.0268	0.0015	0.2070	0.0256	0.0059	0.7460
	Protein yield	237.58	-0.1694	-0.0029	366.79	18.04	0.1875
3	Milk yield	318387.39	6133.82	0.0756	373702.59	23933.87	0.2408
	Fat %	0.0780	0.0070	0.3312	0.0754	0.0059	0.2892
	Fat yield	376.93	-3.98	-0.0427	511.73	36.46	0.2660
	Protein percent	0.0245	0.0037	0.5276	0.0243	0.0025	0.3668
	Protein yield	266,34	2.62	0.0389	321.36	19.91	0.2334
4	Milk yield	265811.80	13013.49	0.1867	376121	35069.63	0.3412

Analysis Number	Trait	Level 1 (low)			Level 2 (high)		
		$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{h}^2$
4	Fat %	0.0700	0.0086	0.4362	0.0755	0.0039	0.1960
	Fat yield	344.53	11.53	0.1295	507.62	37.06	0.2722
	Protein percent	0.0247	0.0055	0.7242	0.0246	0.0040	0.5551
	Protein yield	222.67	8.13	0.1409	316.90	12.91	0.1566
5 (Batch 1)	Milk yield	289091.34	12397.70	0.1645	403097.68	17257.22	0.1642
	Fat %	0.0716	0.0039	0.2044	0.0860	0.0095	0.3970
	Fat yield	413.91	-1.74	-0.0170	536.13	15.51	0.1125
	Protein percent	0.0231	0.0042	0.6106	0.0267	0.0035	0.4592
	Protein yield	271.71	11.49	0.1623	355.50	5.57	0.0617
5 (Batch 2)	Milk yield	289091.34	22560.46	0.2896	33040.65	33040.65	0.3030
	Fat %	0.0716	0.0014	0.0746	0.0860	-0.0035	-0.1708
	Fat yield	413.91	6.71	0.0638	536.13	77.10	0.5029
	Protein percent	0.0231	0.0013	0.2142	0.0267	0.0027	0.3653
	Protein yield	271.71	16.02	0.2227	355.50	53.32	0.5217

Table (3.9)

Summary of the Pooled Analysis of Variance for the Two Sets of Data

	Low		High	
	d.f.	M.S.	d.f.	M.S.
Milk Yield				
Between sires	217	447621.32	217	712004.60
Within sires	6971	274661.58	8624	324745.16
Fat%				
Between sires	217	0.1557	217	0.1650
Within sires	6971	0.0742	8624	0.0802
Fat yield				
Between sires	217	523.65	217	886.56
Within sires	6971	370.44	8624	472.04
Protein %				
Between sires	214	0.0673	214	0.0642
Within sires	6159	0.0250	7362	0.0254
Protein yield				
Between sires	214	360.36	214	533.61
Within sires	6159	247.49	7362	333.62

Table (3.10)

Pooled Estimates of the Between and Within Sire Variance in the Two Sets of Data

	Data set 1				Data set 2			
	Low		High		Low		High	
	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$	$\hat{\sigma}^2_w$	$\hat{\sigma}^2_s$
Milk yield	274661.58	15620.66	324745.16	34497.88	278692.93	13855.59	384214.05	28120.56
Fat percent	0.0742	0.0074	0.0802	0.0076	0.0732	0.0069	0.0797	0.0072
Fat yield	370.44	13.84	472.04	36.93	372.62	11.71	529.80	29.86
Protein percent					0.0250	0.0040	0.0254	0.0039
Protein yield					247.49	10.61	333.62	20.27

the pooled K-values, heritability estimates, and estimates of the genetic correlations between the low and high levels of production in the two data sets. The corresponding weighted values are shown between brackets. It is evident that the weighting did not change the results substantially. All yield traits exhibited a significant rise in heritability from the low to the high level while heritabilities of percentage traits remained unchanged. Estimates of the genetic correlation for all traits were consistently high and ranged from 0.89 in the case of milk yield to 1.12 in the case of protein yield.

Table (3.11)

Estimates of the Genetic Correlation, Heritability and K values from the Weighted and the Pooled Analysis of Variance*

Data Set Number	Trait	$\hat{r}_G$	Low Level		High Level	
			K	$\hat{h}^2$	K	$\hat{h}^2$
1	Milk yield	0.8902 ±0.1276	11.07	0.2152 ±0.0539 (0.2173)	11.23	0.3841 ±0.0648
	Fat percent	0.9633 ±0.0962	11.07	0.3609 ±0.0632 (0.3631)	11.23	0.3441 ±0.0457 (0.3437)
	Fat yield	0.9606 ±0.1784	11.07	0.1440 ±0.0413 01432		0.2902 ±0.0447 (0.2929)
2	Milk yield		10.64	0.1897 ±0.0424	9.87	0.2728 ±0.0443
	Fat percent		10.64	0.3465 ±0.0458	9.87	0.3319 ±0.0455
	Fat yield		10.64	0.1219	9.87	0.2134
	Protein percent	0.9733 ±0.0609	10.64	0.5477 ±0.0488 (0.5453)		0.5366 ±0.0486 (0.5390)
	Protein yield	1.1644 ±0.1742	10.64	0.1644 ±0.0418 (0.1649)		0.2291 ±0.0434 (0.2279)

* Weighted estimates are inserted between brackets.

Section 5: Discussion

Introduction

As may be apparent from the review of literature given in chapter 2 the problem of genotype X environment interactions may be formulated in two questions:

1. Whether the phenotypic expression of genotypes in different environments varies in such a way that the rank of these genotypes is significantly different from one environment to the other.
2. Whether there is a change in the genetic variance. This may or may not be accompanied by a change in the phenotypic variance and consequently the heritability may change from one environment to the other.

In the previous chapters consideration of the problem of genotype X environment interaction was confined to whether it was possible to detect statistically significant interactions or not. Due to the fact that the bulls under study were old selected sires no attempt was made to compute heritabilities or genetic correlations. This statistical approach cannot tell whether the interaction is of any biological importance. The reason why this analysis was initiated was to attempt to answer the first question stated above by obtaining estimates of the genetic correlation and to answer the second question by obtaining estimates of heritability in different levels of production.

The Power of the Analysis

The average effective number of daughters per sire in this analysis is about 11. This should give reasonably low

3.1 The Standard Error for the Genetic Correlation between Performance in two Separate Environments.

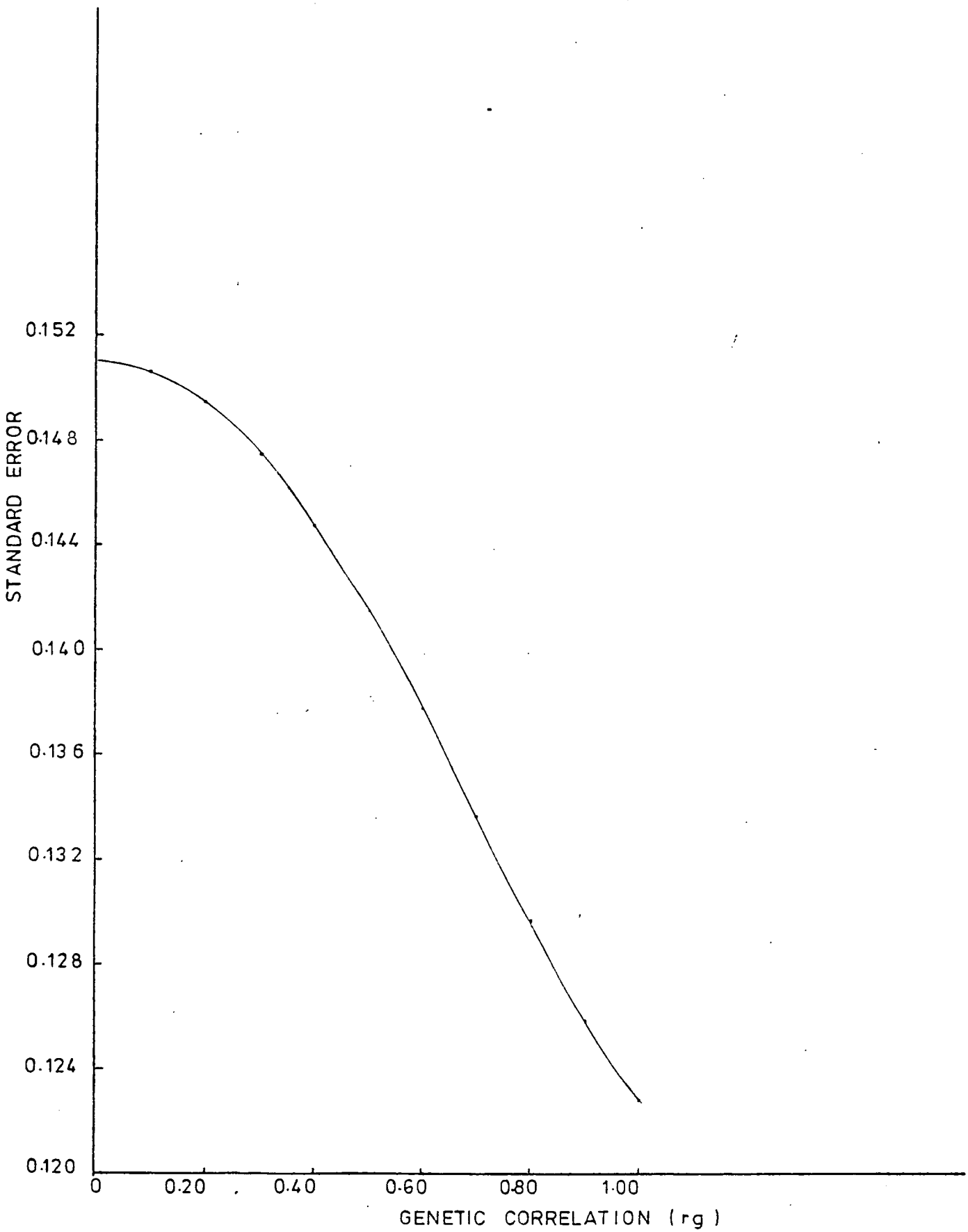


Fig. 3.1

efficient enough for estimating heritabilities but apparently a higher family size is required for estimating the genetic correlation.

Another doubt about the power of this analysis arises from the fact that old proven sires with large numbers of daughters spread over many herd-year-seasons were used to create connections among herd-year-seasons and to facilitate the estimation of the between young sires' mean squares. This was made necessary because of the disconnectedness detected in the data when young sires were analysed separately. Consequently they were nested together with old sires within batches and separate estimates of variance were obtained for each group. The old sires were chosen only on the basis of contemporeneity and the total number of herd-year-seasons in which they were represented. There is a possibility that any interactions detected among young bulls could have been generated because of interactions among old sires or differential use of them in different levels. The unbalanced association of a young sire with an old one over levels could also generate such interactions. The correlations between progeny groups means for old bulls in the low and high levels for all sets of data were found to range between 0.7 and 0.9, roughly similar to that estimated for young sires and since a large number of old bulls was used to dilute the effect of any outlying sire it was concluded that connections through old sires could not have generated interactions among young sires.

The Accuracy of Progeny Testing in Two Environments

The opinion that is is easier to distinguish between genotypes under favourable environmental conditions than under poor conditions has often been expressed in the literature. It is assumed that a favourable environment allows the full realization of the individual's genetic potential while a poor environment suppresses the expression of genetic differences between individuals. To verify this opinion we need to know whether there is more genetic variation relative to the total variance at the higher level of environment and by answering that we can determine whether greater accuracy can be obtained by testing in one environment rather than the other in order to achieve the maximum possible genetic progress. The definition of the accuracy of the progeny test that will be used is that given by Mason and Robertson (1956) as the correlation between future and tested daughters i.e.

$$\text{Accuracy} = r = \sqrt{\frac{\sigma_s^2}{\sigma_s^2 + \sigma_w^2/n}}$$

It is clear that the weighting method used did not make much difference to the estimates of heritability obtained in the two levels for all five traits under study. This is because there were no large differences between sets of data and different levels of production in the value of the coefficient of the components of variance (the average effective number of daughters per sire). As can be seen from table (3.10) the relative magnitude of the between sire components of variance increased significantly for all yield traits with the rise in

the mean level of production. The increase in genetic (sire) variance was significant in the case of milk yield where it increased by 121%. Fat production genetic variance increased by 167% from low to high, and protein production by 91%. The change in the genetic variance of fat % and protein % was insignificant ranging between 1 and 2%. The coefficient of variation for milk production and fat production dropped slightly from 13.8% to 12.2% and from 13.5 to 12.3% respectively. In the case of fat% the coefficient of variation showed a slight rise from 7.2% to 7.4%. The result of this is a rise in the estimate of heritability of milk production by nearly a factor of 2, a similar rise in the estimate of heritability of fat production and a smaller increase in the case of protein production. There was a slight drop in the heritabilities of both fat % and protein %. These results are in agreement with the findings of Mason and Robertson (1956) who analysed data on the progeny of 152 sires with an average number of daughters of 57. The records were made in 1500 herds and the average number of records per herd-year was 5.9. They found an increase in the heritability of milk production from 0.05 in the low level to 0.22 in the high. However, they also estimated the heritability of butterfat percentage to be 0.27, 0.47 and 0.49 in the low, medium and high levels respectively. In the case of fat percentage they did not adjust for environmental differences between herds, assuming that such differences were mainly genetic. If environmental differences were significant and different in importance between the low and high levels,

then this could bias their results significantly. Van Vleck (1963) analysed a total of 85137 first and second lactation records of Holstein cows. He detected a rise in the heritability of milk production from 0.19 to 0.28 and he suggested that for the purpose of analysing dairy cattle data a non-linear model might be more appropriate than the usual additive effects model. Robertson et al (1960) analysed data on a sample of highly selected bulls with large numbers of daughters. Fifty seven Friesian bulls with at least 100 effective daughters and eight Ayrshire bulls with an average of 116.2 effective daughters were analysed. Higher genetic variance within progeny groups was found in the high level herds but since the variance within progeny groups, within herd-years increased to the same extent no rise in heritability was detected and the authors concluded that the accuracy of the progeny test was effectively the same at all management levels.

There is no certainty in the literature as to the causes of differences in yield between high producing and low producing herds. However if these differences are a reflection of a greater adherence to the practice of feeding according to production in the higher levels of management, then this might offer some explanation to the reasons behind the increase in the between sire variation in the high level herds. The practice of feeding according to production should favour the daughters of those sires that are superior from genetic reasons and so the between sire variation will be expected to be greater. A possible reason for the relatively higher within sire variance

in the low levels of management is the lack of environmental uniformity and the poorer control over the environment. The increased possibility of errors in the parentage of heifers could be another source of non-genetic variation. Another possible source of complications is the differential treatment of daughters of the better sires although in this case this is unlikely due to the fact that these are young bulls with small numbers of daughters. The combined effect of all these factors on the estimates of heritability of different traits in the two production levels cannot be predicted with any certainty due to the variable effect they could have on the relative amounts of the between and within sire variance.

These results indicate that for all three yield traits there is more genetic variability relative to the total variance in the high levels of management. In the case of percentage traits, despite the slight changes in the estimates of heritability from the low to the high there is no significant difference of heritability between the two levels of management and consequently the accuracy of progeny testing with regards to these two traits may be assumed to be the same over all herd levels. In the discussion that follows the implications of differences in heritabilities of yield traits over levels will be considered and to start with it will beassumed that the sires are similarly ranked in the two levels i.e. that the genetic correlation is unity.

The differences in heritability between the low and high levels inevitably cause differences in the degree of accuracy of the progeny test. The estimates of accuracy as defined by

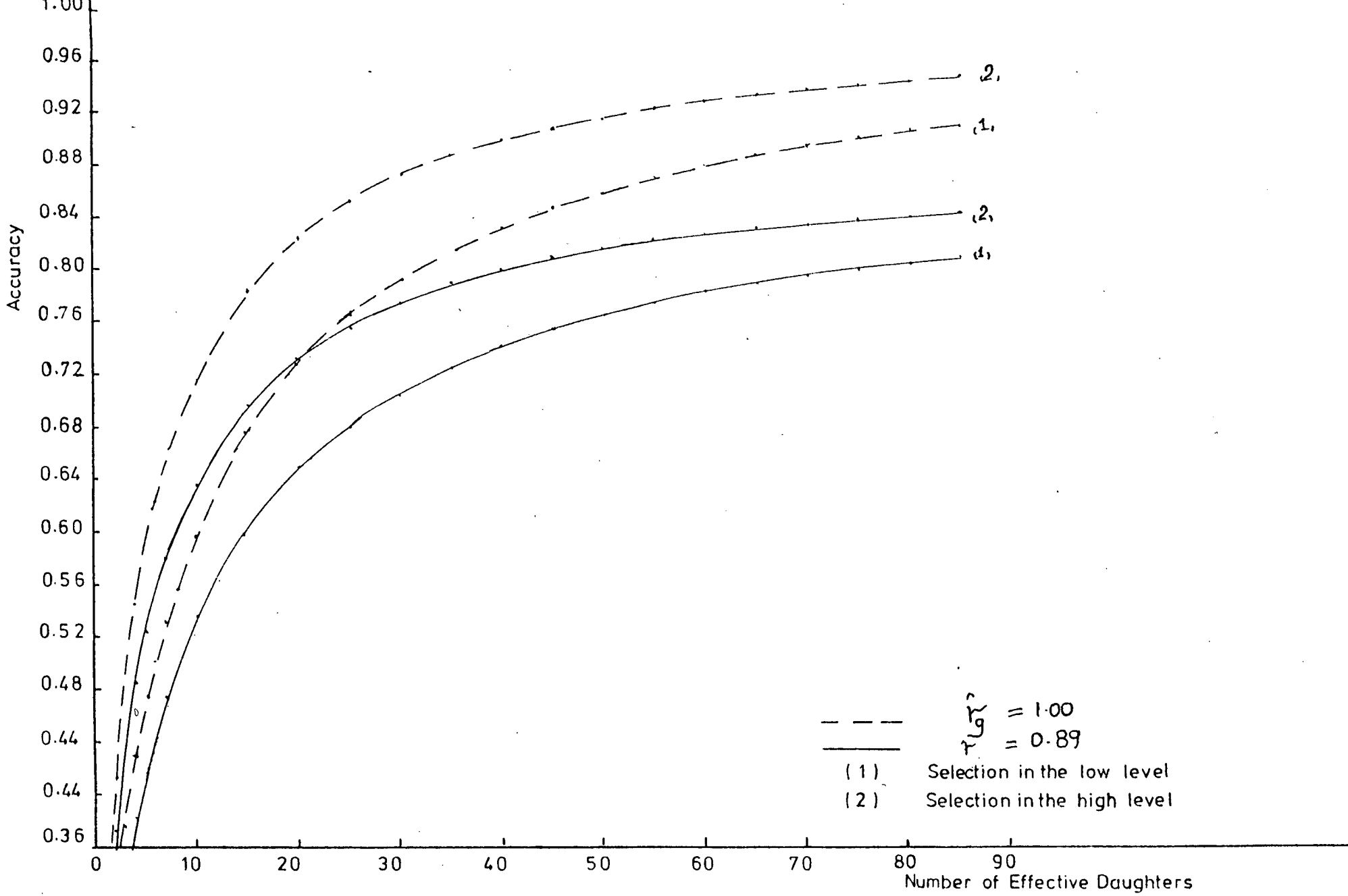
Mason and Robertson (1956) for the three yield traits are shown in table (3.12), assuming that there are equal numbers of

Table (3.12)
Estimates of Accuracy of Progeny Testing in Different Levels

Trait	D.F.	Effective Number	Low		High	
			r	n	r	n
Milk yield	217	11	0.38	(17.58)	0.54	(9.41)
Fat yield	217	11	0.29	(26.77)	0.46	(12.78)
Protein yield	214	10	0.30	(23.34)	0.38	(16.46)

effective daughters in each level which is not far from what is the case in this analysis. The numbers required for a test accuracy of 0.5 (equivalent to $6_w^2/6_s^2$, in the manner of Robertson et al (1960) are inserted between brackets. To achieve an accuracy level of 0.5 it is clear that twice as many daughters are required in the low as in the high level for milk yield and fat yield and close to 1.5 as many daughters are required in the case of protein yield. The values of accuracy for the three traits in the two levels were computed using estimates of heritability obtained from the combined analysis.

Using the formula derived by Mason and Robertson (1956) the accuracy of selection in the high (low) level and making the observations in the low (high) level is plotted against the number of effective daughters in each level in figure (3.2). The number of effective daughters is assumed to be the same in the two levels and the graph is plotted assuming a genetic correlation of 1.00 in one instance and 0.89 in the other. The



3.2 The Effect of the Numbers of Effective Daughters and the Genetic Correlation on Selection for Milk Yield in two Levels of Production.

increased accuracy of selection that can be gained by selecting in the higher level of management due to the higher heritability may be used in one of two ways: By reducing the actual numbers of records that are necessary to test each bull or by decreasing the numbers of bulls being tested. Figure (3.2) illustrates how much saving (in terms of numbers of records) is made by switching selection for milk yield from the low level to the high and rearing the animals in the former. Heritabilities in the low and high levels were assumed to be 0.21 and 0.38 respectively. To achieve a certain level of accuracy of selection the number of records required for selecting the animals in the high level and observing them in the low is on average about half the number required for selecting animals in the low level and rearing them in the high assuming a genetic correlation of unity. For particular numbers of progeny (n) the difference in accuracy that can be attained by selecting in one of the two levels and rearing in the other is greatest at intermediate values of n. Similar arguments to those used apply to the other two yield traits with less advantage being gained, by the way of reductions in the number of records, in the case of protein yield because of the smaller difference in heritability and slightly more advantage in the case of fat yield.

The existence of differences in the amount of genetic variability relative to the total variance indicates the presence of some form of genotype X environment interactions. The results seem to lend support to the view that poor environments tend to obscure genetic differences between individuals and

they seem to suggest that greater potential genetic progress can be achieved by concentrating testing and selection in the high production herds. The actual gains made by adopting such a policy will in reality be smaller than anticipated above since the present practice is to test bulls in both levels of management rather than in one or the other as was assumed hitherto. If testing is concentrated in the high management herds with little or no testing in the low level herds, this will inevitably reduce either the number of bulls that can be tested or the number of records on which each bull is tested. This reduction will have to be balanced against the gains procured by concentrating testing and selection in the better herds. This practice is also likely to impose a genetic load on the better herds since the majority of the heifers in a testing programme will be daughters of inferior bulls that will eventually be discarded, although it may be argued that those young bulls are more often produced in the better herds and they are on average likely to be better than their fathers. Whether this genetic load is real or imaginery the resistance of herd managers to the testing of young bulls is real and is an obstacle to be faced. The above applies to the general testing programme; however, if a policy of contract mating for young sires is being followed, as is apparently the case with some MMB bulls, it would certainly seem advisable to choose herds for this purpose from among the higher management herds.

The Correlation Between Genotypic Expressions in Different Environments

We now consider the situation where the rank of genotypes actually changes from one environment to the other i.e. where the genetic correlation coefficient is significantly less than unity. The implications to the breeder in the event of the existence of this type of interactions are clear; a different strain or breed of animals has to be developed for each environment capable of supporting one. In this sense this form of genotype X environment interactions is of far greater practical importance than the form that has been discussed previously.

In this analysis the genotypic correlations were computed according to Falconer (1960) and they were reasonably high for all traits and their differences from unity may be attributed to sampling errors. However the true parameters may well be lower than the estimates given for the same reason. As the estimates were fairly high, the estimate of the genetic correlation between performance in the two environments for milk yield will be taken as an example to illustrate the effect of a correlation, that is different from unity, on the arguments about the accuracy of selection and the numbers required in each environment to achieve a given accuracy.

Figure (3.2) illustrates the effect of a true correlation of 0.89. The effect of a real correlation of that magnitude is to significantly reduce the accuracy of selection that can be attained by any particular number of effective daughters compared to the situation in which the genetic correlation was

assumed to be one. The difference in the accuracy of selection between the situation where $\hat{r}_g$ is one and where $\hat{r}_g = 0.89$ is small at low effective numbers and increases rapidly with the increase in numbers and becomes constant as the former draws to a limit of 1 and the latter to a limit of 0.89.

The estimates of the genetic correlation for all five traits between the two levels of environment suggest that the ranking of sires will be the same in both environments irrespective of where their daughters were produced. This is in agreement with the findings of Robertson et al (1960) who found a range of correlations for milk yield between 0.87 and 0.96 among Friesian bulls and from 0.83 to 0.91 in Ayrshire bulls; the data on the latter being smaller. They suggested that the true correlations are likely to be higher due to the inflation of the within sire variance by errors in the estimation of the breeding value of bulls. Van Vleck (1963) also found no evidence of change of rank among four different levels of herd-mate production in a rather large set of data described above. In fact it seems most reports agree that the genetic correlations over different production levels - despite the large differences in the definitions of such levels - are high, apart from Brumby (1961) who reported a genetic correlation of 0.11 among two sets of identical twins. It seems likely, however, that the environmental differences in this latter study are much higher than those previously mentioned in other studies.

Conclusions

It is clear from the literature that the magnitude of interaction varies with the width of the genetic base of the population under study and the range of environments to which it is exposed. Dairy cattle have been subjected to continuous selection within breeds and the maintainence of the 'purity' of the breeds has been one of the major concerns of breed societies. The extensive use of artificial inseminatin in dairy cattle production has probably contributed further to the narrowing of the genetic width of dairy breeds. The range of environments to which dairy cattle are exposed is limited by the very nature of dairy production in which individual attention is necessary to keep a profitable business in a competitive industry. The effect of these two factors has been reflected in the fact that many investigators have reported either finding negligible or no genotype X environment interactions in the usual sense i.e. changes in the ranking of sires over environments. This analysis within its limitations that were set out earlier indicates that no such changes in ranking occur over the two levels of production studied. The changes that have occurred in the dairy industry since the study of Robertson et al (1960) have not changed the validity of their essential conclusion: that sires may be tested in any management level and that no special stains or breeds of animals need to be developed for different management levels.

This study also confirms the view that in dairy production yield traits exhibit higher heritabilities in the higher management levels. At first sight this may seem to suggest that

note error in page ~~numbering~~ order

CHAPTER 4

General discussion and Conclusions

It is important here to bring together the findings of the two analyses undertaken within this project and relate them to the stated aims of the study. The analysis of daughters of widespread use bulls described in chapter (2) shows that the component of variance for the sire X level of production interaction is likely to be small. Product moment correlations obtained for all classifications of herds ranged between 0.8 and 0.9 for milk yield. However, a few sires deviated substantially from the general regression line as can be seen from figures (2.1 and 2.2). As these bulls are heavily used it may be worthwhile to devise a simple test to detect those few that exhibit large interactions and set up a system to make the best use of them. It is likely that such a test will involve the division of herds into production levels and obtaining separate estimates for each bull. Such a test, if carried out on young bulls, will probably be useless because of the small numbers of daughters involved.

Despite the fact that the analysis on the daughters of widespread use bulls showed no interactions in the usual sense i.e. changes in ranking, the results indicated the existence of important shifts in 'variance' from one level of production to the other. Since the data consisted entirely of records of daughters of highly selected sires, it was considered important to investigate further those shifts in variance using a set of young test bulls. The estimates of the genetic correlation obtained from this analysis were very high for all five traits and it confirmed the finding from the first analysis that sires

rank similarly over different production levels. However, the family size in this analysis was rather small and this reduced significantly the power of the analysis to detect important differences from unity in the genetic correlation.

In a situation where the genetic correlation is one, the only other factor affecting the accuracy of progeny testing at constant numbers of progeny in each herd level is the heritability. The analysis on test bulls daughters indicates that there is a significant difference in heritability between the two levels of production (at the 5% level). This causes important differences in the numbers of effective daughters required for testing in each level to obtain the same accuracy. This can be taken account of in one of the following ways:

a) By testing exclusively in the level where the heritability is higher or by concentrating most of the testing in it. This will inevitably result in a reduction of the total number of cows available for testing each bull or a reduction in the total number of bulls that can be tested. This will have to be weighed against the increased dividends that accrue from testing in a high heritability environment. The analysis on young bulls daughters shows that in the case of milk yield the heritability in the low level is 0.21 and in the high level it is 0.38 giving an accuracy ratio of the low to the high level of 1.89. This means that, for example, at an accuracy of 0.92 in both levels of production, 50 effective daughters at the high level will be equivalent to 95 effective daughters at the low level. This is in accordance with the indications of a higher sire variance at the higher level of production obtained from the analysis on widespread use sires described in chapter (2).

selection should be carried out in the better herds, and as the genetic correlation may for all practical purposes be taken to be unity, their ranking will be expected to be the same in the lower management levels. However the economies that can be brought about in this way have to be looked at in comparison with the costs incurred by the loss in potential testing capacity when low level herds are excluded. This analysis shows that the saving in the number of records that can be made by testing exclusively in the high level is about half. In practice this is likely to be lower since already part of the testing is done in the high level but it must be worth investigating further the differences in heritability between management levels and with more precise knowledge of the present ratios of testing over different management levels it should be possible to determine whether it is worthwhile to lay more emphasis on testing in the better management herds. In the case of contract mating bulls there should be no problem in concentrating testing in the high management level.

b) By scaling the actual observations on heifers. Assuming that the genetic correlation is one the only important element determining the scaling factors is the heritability.

c) By obtaining separate estimates for each sire in each level of production and combining these two estimates using an index that is determined by the heritability, the genetic correlation and the appropriate ratios of animals in each environment. The equation used to calculate such an index for this analysis is the following:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \sigma^2_{s1} + \frac{\sigma^2_{w1}}{n_1} & Cov_B \\ Cov_B & \sigma^2_{s2} + \frac{\sigma^2_{w2}}{n_2} \end{pmatrix}^{-1} \begin{pmatrix} \sigma^2_{s1} & Cov_B \\ Cov_B & \sigma^2_{s2} \end{pmatrix} \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \qquad (5.1)$$

Where b_1 and b_2 are the weights in the index to be attached to sire estimates from the low and the high level of production respectively, n_1 and n_2 are the respective numbers of effective daughters in each environment and R_1 and R_2 are the proportions of animals reared in each environment. Assuming that the effective numbers of daughters are equal in the two environments, the index weights for milk yield were calculated and are shown in table (5.1), assuming that in one instance $n_1 = n_2 = 11$ and in the other $n_1 = n_2 = 22$.

Table (5.1)

Weights for Combining Estimates of Sires from the Low and High Level of Production

(1) $n_1 = n_2 = 11$

$\hat{r}_g$		Rearing Environment: Low ($R_1 = 1, R_2 = 0$)	High ($R_1 = 0, R_1 = 1$)	½ low + ½ high ($R_1 = \frac{1}{2}, R_2 = \frac{1}{2}$)
0.89	b_1	0.26	0.28	0.27
	b_2	0.24	0.45	0.34
1.00	b_1	0.22	0.33	0.28
	b_2	0.28	0.42	0.35

(2) $n_1 = n_2 = 22$

0.89	b_1	0.36	0.32	0.34
	b_2	0.27	0.57	0.42
1.00	b_1	0.27	0.41	0.34
	b_2	0.34	0.51	0.43

The findings of both analyses are in accordance with previous evidence with regard to both the problem of changes in rank and heritability estimates in different levels of production. There is only one previous analysis on British data by Robertson (1960) which showed a high correlation between different levels of production but failed to detect significant changes in heritability. However, there is ample evidence from American and other data indicating important changes in heritability over levels e.g. Mason and Robertson (1956) and Van Vleck (1963).

The increase in heritability in the higher level of production may have important implications to the problem of devising indices for the selection of bull mothers. It also means that in a situation where contract mating is considered the breeder should

endeavour to contract only the high production herds for that purpose.

The findings from the analysis on wide-spread use bulls indicate that although the interactions of age effects with sires and with some classifications of the herd environment are statistically significant, they are not important enough to warrant specific action to account for them. However, curvilinearity in the correction of milk yield for age effects was found to be unimportant except in the low producing spring herds. It is also clear that with regard to age X level interactions the season of herd calving is the most significant classification in the sence that it is the classification that showed the highest age X level interactions.

It is important to note that the classification of the herd environment in both analysis was made on the basis of the heifers' mean milk yield and that it was not related to either fat or protein percentages or to any concrete measure of management level. In any future study it may be more informative to base the classification of herds on actual management practices rather than the herd average yield.

We may now usefully summarise the findings of the two analyses:

1. Sires are ranking similarly in different production levels and hence there is no need to develop special strain for each level. However, there may be a case for analysing data of wide-spread use sires, after they accumulate sufficient numbers of daughters, in separate production levels in order to detect the few sires that deviate from the general rule and make the best use of them.

2. Taking the findings of the two analyses and previous evidence on the subject, it is likely that there is an increase in heritability with the increase in the mean level of herd production. It is difficult to specify a biological reason for such an increase but it is possible that the better control of environment, greater adherence to feeding according to production and better individual attention in the high level herds may have contributed to this increase in heritability.

REFERENCES

Allaire, F.R. and Gaunt, S.N. 1965. First lactation contemporary comparisons as indicators of environmental influences on daughter records used for sire evaluation. J. Dairy Sci. 48:

Barker, J.S.F. and Robertson, Alan, 1966. Genetic and phenotypic parameters for the first three lactations in Friesian cows. Anim. Prod. 8: 221-240.

Bereskin, B. and Touchberry, R.W. 1966. Some relationships of body weight and age with first lactation yield. J. Dairy Sci. 49: 869-873.

Bowman, J.C. 1972. Genotype X environment interactions. Annales de Genetique et de selection Animale. 4 (1): 117-123.

Brumby, P.J. 1961. The causes of differences in production between dairy herds. Anim. Prod. 3: 277-

Burdick, J.M. and McGilliard, L.D. 1963. Interaction between sires in artificial insemination and management of dairy herds. J. Dairy Sci. 46: 452-458.

Burnside, E.B. and Rennie, J.C. 1961. The heritability of milk yield at different levels of production and the effect of production differences on dairy sire appraisals. J. Dairy Sci. 44: 1189-

Burnside, E.B. and Rennie, J.C. 1968. Sire evaluation and herd level of milk production. Can. J. Anim. Sci. 48: 97-102.

Dickerson, G.E. 1962. Implications of genetic-environmental interaction in animal breeding. Anim. Prod. 4: 47-

Falconer, D.S. 1952. The problem of environment and selection. Am. Naturalist, 86: 293-

Falconer, D.S. 1960. Introduction to quantitative genetics. Oliver and Boyd. Edinburgh and London.

Fairchild, T.P., Corley, E.L. and Tyler, W.J. 1963. Evaluating dairy sires by herd-mate comparisons. J. of Anim. Sci. 22: 1109-

Freeman, A.E. 1969. Estimates of a genotype by environmental interaction and effects of contemporaneity upon growth and production from cattle twins. J. Dairy Sci. 52: 926-

Freeman, A.E. 1973. A quarter century of artificial insemination of dairy cattle are changes and new approaches indicated? J. of Anim. Sci. 37 (3): 658-667.

Gacula, M.C., Gaunt, S.N. and Damon, R.A. 1968. Genetic and environmental parameters of milk constituents for five breeds. I. Effects of herd, year, season and age of cow. J. Dairy Sci. : 428-437.

Gravir, K. and Hickman, C.G. 1966. Importance of lactation numbers, age and season of calving for dairy cattle breed improvement. Can. Dep. Agr. Pub. 1239.

Guiard, V. and Herrendorfer, G. 1977. Estimation of the genetic correlation coefficient by half sib analysis if the chacters are measured on different offsprings. Biom. J. 19 (1): 31-36.

Hammond, J. 1947. Animal breeding in relation to nutrition and environmental conditions. Biol. Rev. 22: 195-213.

Hancock, J. 1953. Studies in monozygotic cattle twins. VII. The relative importance of inheritance and environment in the production of dairy cattle. N.Z.J. Sci. Tech., Agric. 35: 67-92.

Hargrove, G.L. 1974. Rate of maturity of dairy females. J. Dairy Sci. 57: 328-

Hartman, O. 1968. Contribution to the problem of selection in different environmental levels. Atti. II Simp. Int. Zootec., Milan, 1967: 191-193.

Harvey, W.R. 19 . Least-squares analysis of data with unequal subclass numbers. U.S. Dep. of Agric.

Harville, D.A. and Henderson, C.R. 1966. Interrelationships among age, body weight and production traits during first lactations of dairy cattle. J. Dairy Sci. 49: 1245.

Hickman, C.G. 1962. Effects of level of herd environment. I. Relationship between yield and age. J. Dairy Sci. 45: 861-864.

Hickman, C.G. and Gravir, K. 1968. Yield adjustment for differences in age at calving. Acta. Agric. Scan. 18: 199.

Hickman, C.G. and Henderson, C.R. 1955. Components of relationship between level of production and rate of maturity in dairy cattle. J. Dairy Sci. 38:

Hickman, C.G. 1957. Production studies on experimental farms dairy herds. I. Age correction of 180-day milk yield. Can. J. Anim. Sci. 37: 123-129.

Hickman, C.G. and Henderson, C.R. 1955. Components of the relationship between level of production and rate of maturity in dairy cattle. J. Dairy Sci. 38: 883.

Hillers, J.K. and Freeman, A.E. 1965. Differences between sires in rate of maturity of their daughters. J. Dairy Sci. 48: 1680 -

Kelleher, D.J., Freeman, A.E. and Lush, J.L. 1967. Importance of bull X herd-year-season interaction in milk production. J. Dairy Sci. 50: 1703-1707.

Lee, A.J. 1972. Month, year and herd effects on age adjustment of first lactation milk yield. J. Dairy Sci. 56: 683.

Lee, A.J. 1974. Month, year and herd effects on age adjustment of first lactation milk yield. J. Dairy Sci. 56: 332.

Lee, A.J. and Hickman, C.G. 1972. Age and herd adjustment of first lactation milk yield. J. Dairy Sci. 55: 432.

Legates, J.E. 1962. Heritability of fat yields in herds with different production levels. J. Dairy Sci. 45: 990-993.

Legates, J.E., Verlinden, F.J. and Kendrick, J.F. 1956. Sire-by-herd interaction in production traits in dairy cattle. J. Dairy Sci. 39: 1055.

Lush, J.L. 1945. Animal Breeding plans. Ames: Iowa State College Press, 3rd ed.

Lytton, V.H. and Legates, J.E. 1966. Sire by region interaction for production traits in dairy cattle. J. Dairy Sci. 49: 874-878.

Mao, I.L. and Burnside, E.B. 1969. Sire by herd environment interaction for milk production. J. Dairy Sci. 52: 1055-1062.

Mason, I.L. and Robertson, Alan. 1956. The progeny testing of dairy bulls at different levels of production. J. Agric. Sci. 47: 367-375.

McDaniel, B.T. and Corley, E.L. 1966. Environmental influences on age correction factors. J. Dairy Sci. 49: 736-

Miller, R.H. 1964. Biases in the estimation of the regression of milk production on age. J. Dairy Sci. 47: 855-

Miller, P.D., Lentz, W.E. and Henderson, C.R. 1970. Joint influence of month and age of calving on milk yield of Holstei cows in the Northeastern United States. J. Dairy Sci. 53: 351-

Nelder, J.A. and Wedderburn, R.W.M. 1972. Generalized linear models. J. Royal Stat. Soc. 135 (A): 370-384.

Pani, S.N. and Lasley, J.F. 1972. Genotype X environment interactions in animals. Theoretical considerations and review findings. Res. Bull. 992. Univ. of Missouri-Columbia Coll. Agric. Exp. Station.

Pirchner, F. 1969. Population Genetics in animal breeding. San Francisco, Calif.: W.H. Freeman and Co.

Richardson, D.O., Owen, J.R., Plowman, R.D., and Miles, J.T. 1971. Importance of sire X ration interactions in production and feed intake traits of dairy cattle. J. Dairy Sci. 54: 1518-1525.

Robertson, Alan. 1959. The sampling variance of the genetic correlation coefficient. Biometrics. 15: 469-485.

Robertson, Alan, O'Connor, L.K. and Edwards, J. 1960. Progeny testing dairy bulls at different management levels. Anim. Prod. 2: 141-152.

Robertson, Alan., Stewart, A. and Ashton, E.D. 1956. The progeny assessment of dairy sires for milk: the use of contemporary comparisons. Proc. Brit. Soc. Anim. Prod. 43:

Sargent, F.D., Butcher, K.R. and Legates, J.E. 1967. Environmental influences on milk constituents. J. Dairy Sci. 50: 177-184.

Searle, S.R. 1962. Age and herd effects in New Zealand dairy cow records. J. Dairy Sci. 45: 82-

Searle, S.R. 1965. The value of indirect selection. I. Mass selection. Biometrics 21: 682-707.

Searle, S.R. 1978. The value of indirect selection. II Progeny Testing. Theoretical and Applied Genetics. 51:

Searle, S.R. and Henderson, C.R. 1959. Establishing age correction factors related to the level of herd production. J. Dairy Sci. 42: 824-

Sprecht, L.W. and McGilliard, L.D. 1960. Rates of improvement by progeny testing in dairy herds of various sizes. J. Dairy Sci. 43: 63-75.

Syrstad, Ola. 1965. Studies on dairy herd records. II. Effect of age and season of calving Acta Agric. Scand. 15: 31-64.

Syrstad, Ola. 1977. The interaction between genotype and level of concentrate feeding in milk production.

The Milk Marketing Board of England and Wales. The report of the Breeding and Production Organisation, 1973/74. Thames Ditton, Surrey, No. 24.

The Milk Marketing Board of England and Wales. Dairy Herd Census, 1970.

Tong, A.K.W., Kennedy, B.W. and Moxley, J.E. 1977. Sire by herd interactions for milk yield and composition traits. Can. J. Anim. Sci. 57: 383-388.

Touchberry, R.W., Rottensten, K., and Anderson, H. 1960. A comparison of dairy sire progeny tests made at special Danish testing stations with tests made in farmer herds. J. Dairy Sci. 43: 529-

Van Vleck, L.D. 1963. Genotype and environment in sire evaluation. J. Dairy Sci. 46: 983-987.

Van Vleck, L.D. 1966. Change in variance components associated with time and increase in mean production. J. Dairy Sci. 49:

Van Vleck, I.D. and Henderson, C.R. 1961. Variance and covariance components in part lactation milk and fat records. J. Dairy Sci. 44: 1870-1876.

Waddell, L.H. and McGilliard, L.D. 1959. Influence of artificial breeding on production in Michigan dairy herds. J. Dairy Sci. 42:

Wright, S. 1939. Genetic principles governing the rate of progress of livestock breeding. Proc. Amer. Soc. Anim. Prod. 32nd Ann. Meet.: 18-26.

Wunder, W.W. and McGilliard, L.D. 1967. Seasons of calving and their interactions with age for lactational milk yield (Abstr.) J. Dairy Sci., 50: 986-987.

Yamada, Y. 1962. Genotype by environment interaction and genetic correlation of the same trait under different environments. Jap. J. Genet. 37: 498-509.

Acknowledgement

I wish to express my thanks to the Milk Marketing Board of England and Wales for providing the data used in this project. My thanks also to the British Council and the University of Khartoum for awarding me the scholarship. I also wish to express my gratitude to my supervisors; W.G. Hill, C.J.M. Hinks and R. Thompson for their valuable advice and support during the writing of this thesis. Finally my thanks to my family for their continued understanding and encouragement.